

Residue Identity = 100%	Matches = 0	Conservative Substitutions = 349	Mismatches = 0				
Gaps =							
X	10	20	30	40	50	60	70
MTAASMGPRVAVFVLLALCSRPAVGONCGSGPCRPDEPAPRCPAGVSLVLDGCGCCRCVCAKQIGELCTERO							
MTAASMGPRVAVFVLLALCSRPAVGONCGSGPCRPDEPAPRCPAGVSLVLDGCGCCRCVCAKQIGELCTERO							
X	10	20	30	40	50	60	70
PCDPKHGILFDGSGPANRKIGVCTAKDGAPICIFGGTVYBSGFSFOSSCKYOCCTCLDGAVGCMPLCSMDVRLP							
PCDPKHGILFDGSGPANRKIGVCTAKDGAPICIFGGTVYBSGFSFOSSCKYOCCTCLDGAVGCMPLCSMDVRLP							
X	80	90	100	110	120	130	140
SPDCFPFPRVKLPGKCEEWVCDPEKDQTVVGPALAAVLEDTFGPDTMIRANCLVQTTEWSACSKTCGMG							
SPDCFPFPRVKLPGKCEEWVCDPEKDQTVVGPALAAVLEDTFGPDTMIRANCLVQTTEWSACSKTCGMG							
X	150	160	170	180	190	200	210
ISTRVTNDNASCRLEKOSRLCMVRPCEADLEENIKKKKKCIPTPKISKPIKFELSGCTSMKTYRAKFGCVCT							
ISTRVTNDNASCRLEKOSRLCMVRPCEADLEENIKKKKKCIPTPKISKPIKFELSGCTSMKTYRAKFGCVCT							
X	220	230	240	250	260	270	280
DGRCTTPHRTTLPVEFKCPDGEVMMKKNMFIKTACHYNCFGDNDIFESLYRKMVGDMA							
DGRCTTPHRTTLPVEFKCPDGEVMMKKNMFIKTACHYNCFGDNDIFESLYRKMVGDMA							
X	290	300	310	320	330	340	X
DGRCTTPHRTTLPVEFKCPDGEVMMKKNMFIKTACHYNCFGDNDIFESLYRKMVGDMA							
DGRCTTPHRTTLPVEFKCPDGEVMMKKNMFIKTACHYNCFGDNDIFESLYRKMVGDMA							
X	290	300	310	320	330	340	X

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPearch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Wed Sep 17 09:35:08 1997; MagPar time 7.87 Seconds
507.112 Million cell updates/sec
Tabular output not generated.

```

Title:
Description:
Perfect Score:
Sequence:
>US-08-167-628-2
(1-349) from 5408040.pap
2713
1 MTAASMGPPVRVAFVLLALC...
```

Scoring table:

Gap 11

THE

Searched: 96640 seqs, 11439865 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq27

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20

Statistics: Mean 32.775; Variance 141.735; scale 0.231

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2713	100.0	349	20	W11302	Connective tissue gro	3.93e-253
2	2713	100.0	349	16	R79964	Connective tissue gro	3.93e-253
3	2321	92.9	348	5	R23566	Beta-IG-M2.	4.63e-234
4	1460	53.8	351	6	R31599	Chicken nov protein.	4.62e-129
5	1113	41.0	379	5	R25356	Beta-IG-M1.	5.50e-95
6	1002	36.9	375	16	R90919	Connective tissue gro	3.89e-84
7	644	23.7	205	6	R31608	Homologous to chicken	1.93e-49

8	579	21.3	72	6	R31610	Fragment XXVI with ho	3.25e-4
9	430	15.8	71	6	R31609	Chicken nov protein f	4.30e-29
10	412	15.2	84	6	R31602	Encoded by chicken no	2.09e-27
11	405	14.9	84	6	R31603	Polypeptide X homolog	9.43e-27
12	395	14.6	124	10	R46078	CYR61 like protein.	8.10e-26
13	314	11.6	72	6	R31605	Homologous to chicken	2.49e-18
14	312	11.5	70	6	R31604	Chicken nov protein f	3.79e-18
15	278	10.2	75	6	R31601	Chicken nov protein f	4.66e-15
16	268	9.9	76	6	R31600	Chicken nov protein f	3.72e-14
17	157	5.8	22	6	R31612	Fragment XXXI homolog	1.68e-04
18	137	5.1	22	6	R31611	Fragment XXX encoded	3.99e-03
19	138	5.1	271	5	R26994	Rat IGFBP-5.	5.98e-02
20	125	4.6	184	19	R98994	Vascular IGF-1-like gro	9.29e-02
21	123	4.5	258	4	R22253	Sequence of insulin-1	9.29e-02
22	123	4.5	258	4	R22568	Sequence of insulin-1	9.29e-02
23	121	4.5	455	19	W00231	Drosophila morphogeni	1.33e-06
24	121	4.5	455	9	R47261	Pre-pro 60A.	1.33e-06
25	121	4.5	455	9	R60966	Drosophila 60A morpho	1.33e-06
26	121	4.5	455	10	R57981	Drosophila 60A morpho	1.33e-01
27	121	4.5	455	7	R33413	Drosophila 60A morpho	1.33e-01
28	121	4.5	455	7	R33935	Morphogen 60A full le	1.33e-01
29	121	4.5	455	9	R47298	60A.	1.33e-01
30	121	4.5	455	10	R50208	Drosophila 60A morpho	1.33e-01
31	121	4.5	455	10	R60962	Drosophila 60A morpho	1.33e-01
32	121	4.5	455	9	R46745	Drosophila sp. 60(A)	1.33e-01
33	117	4.3	282	13	R79101	Prostaglandin I2 (PGI	2.72e-01
34	117	4.3	282	13	R79102	Prostaglandin I2 (PGI	2.72e-01
35	113	4.2	193	3	P60463	Sequence of c-termimu	5.50e-01
36	113	4.2	2813	3	P60462	Sequence of human von	5.50e-01
37	112	4.1	2813	3	P60053	Sequence of von Wille	6.56e-01
38	109	4.0	1218	19	W05933	Human Serrate-1 (HSL)	1.11e+00
39	105	3.9	272	19	R95329	Insulin-like growth f	2.21e+00
40	105	3.9	272	5	R26995	Human IGFBP-5.	2.21e+00
41	105	3.9	272	5	R25700	IGFBP6.	2.21e+00
42	105	3.9	272	10	R55084	Human insulin-like gr	2.21e+00
43	106	3.9	807	8	R44241	F-spondin (FF5-9).	1.86e+00
44	103	3.8	328	1	R40908	EcoRI-EcoRI fragment	3.11e+00
45	103	3.8	5317	R83016	Recombinant papilla	3.11e+00	

ALIGNMENTS

RESULT	1	
AD	W11302 standard; Protein; 349 AA.	
ID	W11302;	
CD	W11302;	
DE	18-MAR-1997 (first entry)	
DE	Connective tissue growth factor;	
DE	Connective tissue growth factor;	
DE	Proliferative disease; platelet-derived	
KW	tissue growth; repair; umbilical	
KW	tissue growth; repair; umbilical	
KW	antibody; wound healing; cancer;	
KW	inhibitor; protease degradation;	
KW	home sapiens.	
OS	US36324056.	
PN	US36324056.	
PD	10/DEC/1996.	
PF	30-AUG-1991; 732427.	
PP	30-AUG-1991; US-752427.	
PR	30-AUG-1991; US-167628.	
PR	14-DEC-1993; US-167628.	
PR	10-FEB-1995; US-386680.	
PR	(UYSF-) UNIV SOUTH FLORIDA.	

PI Bradham DM, Grotendorst GR;
DR WPI; 97-051180/05.
DR N-PSDB; T51234.
PT New nucleic acid encoding connective tissue growth factor - useful
PT for accelerating wound healing, also for diagnosis and treatment of
PT proliferative disease
PS Claim 9; Column 15-18; 11pp; English.
CC This sequence represents the human connective tissue growth factor
CC (CTGF). CTGF is related immunologically and biologically to
CC platelet-derived growth factor (PDGF), but is encoded by an unrelated
CC gene. CTGF is thought to play a significant role in the normal
CC development, growth, and repair of human tissue, similarly to PDGF. The
CC cDNA encoding this sequence was isolated by screening a cDNA library from
CC human umbilical vein endothelial (HUVE) cells with anti-PDGF antibodies.
CC CTGF can be used to accelerate wound healing. Also, elevated levels of
CC CTGF may be diagnostic of proliferative diseases involving outgrowth of
CC connective tissue cells, such as cancer, fibrotic disease and
CC atherosclerosis. All of these diseases can be treated with reagents
CC reactive with CTGF, such as antibodies (which can also serve as assay
CC reagents). Antisense nucleic acids, and ribozymes could also be used to
CC inhibit CTGF production. The advantage with using CTGF is that it is
CC more stable, and less susceptible to protease degradation than PDGF, and
CC other growth factors involved in wound healing. This is believed to be
CC due to the high Cys content.
SQ Sequence 349 AA;

Query Match 100.0%; Score 2713; DB 20; Length 349;

Best Local Similarity 100.0%; Pred. No. 3.93e-253;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mtaasngpvrvaavfvlalcsrpavqncsgpcrcpdepaprcpagvslvldgcgcrcvc 60
QY 1 MTAASMGpvrvaavfvlalcsrpavqncsgpcrcpdepaprcpagvslvldgcgcrcvc 60

Db 61 akqlgelcterdpdpkhgldfsgspanrkigvctakdgapciffgtvyrsgesfqscc 120
QY 61 AKQlGELCTERDPDPKHGLDFSGSPANRKIGVCTAKDGPAPCFIFGTVYRSGESFQSSC 120

Db 121 kyqctclldgavgcmlcsmdvrlpsdpfprvrklpgkceewvcdpkdqtvvvgpala 180
QY 121 KYQCTCLDgavgcmlcsmdvrlpsdpfprvrklpgkceewvcdpkdqtvvvgpala 180

Db 181 ayrledtfgpdpmtiranclvqtewsacsktcgmigistrvndnasrlekqslcmvr 240
QY 181 AYRL EDTFGPDPTMIRANCLVQTWEWSACSKTCGMIGISTRVNDNASCRLEKQSLCMVR 240

Db 241 pceadleenikkkgkcirtpkiskpikfelsgctsmktyrakfcgvcctdgrcctphrttt 300
QY 241 PCEADLEENIKKKGKCI RTPKISKPIKFELSGCTSMKTYRAKFCGVCCTDGRCTPHRTTT 300

Db 301 lpvefkcpdgevmkmmfiktacachncpgndndifeslyyrkmygdma 349
QY 301 LPVEFKCPDGEVMKMMFIKTCACHYNCPCGNDNDFESLYYRKMYGDMA 349

RESULT 2
ID R79964 standard; Protein; 349 AA.

AC R79964;

DT 12-JUN-1996 (first entry)

DE Connective tissue growth factor.

KW Connective tissue growth factor; CTGF; wound healing; vulnary;
KW cell proliferation; cancer; fibrosis; atherosclerosis; diagnosis
KW therapy; mitogen.
OS Homo sapiens.

FH Key Location/Qualifiers

FT Modified site 28
FT /label= N-glycosylation_site
FT Modified site 225
FT /label= N-glycosylation_site
PN US5408040-A.
PD 18-APR-1995.
PF 30-AUG-1991; 752427.
PR 30-AUG-1991; US-752427.
PR 14-DEC-1993; US-167628.
PA (UYSF-) UNIV SOUTH FLORIDA.
PI Bradham DM, Grotendorst GR;
DR WPI; 95-161147/21.
DR N-PSDB; T04226.

ET New connective tissue growth factor - used to develop prods. for
ET wound healing and for diagnosis and therapy of cell proliferative
ET disorders.

PS Claim 1; Column 19-20; 12pp; English.

CC Novel human connective tissue growth factor (CTGF) (R79964)
CC is related immunologically and biologically to platelet-derived
CC growth factor (PDGF), but is the product of a distinct gene.
CC CTGF is mitogenic and also a chemotactic agent for cells. It is
CC produced by endothelial and fibroblastic cells, and probably acts
CC as a growth factor in wound healing. Recombinant CTGF can be obtd.
CC by expression of cDNA clone DB60R32 (T04226) in transformed host
CC cells. It is used to accelerate wound healing, and to raise
CC antibodies useful in detecting disorders associated with overgrowth
CC of cells, such as cancer, fibrotic diseases and atherosclerosis.

SQ Sequence 349 AA;

Query Match 100.0%; Score 2713; DB 16; Length 349;

Best Local Similarity 100.0%; Pred. No. 3.93e-253;

Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mtaasngpvrvaavfvlalcsrpavqncsgpcrcpdepaprcpagvslvldgcgcrcvc 60
QY 1 MTAASMGpvrvaavfvlalcsrpavqncsgpcrcpdepaprcpagvslvldgcgcrcvc 60

Db 61 akqlgelcterdpdpkhgldfsgspanrkigvctakdgapciffgtvyrsgesfqscc 120
QY 61 AKQlGELCTERDPDPKHGLDFSGSPANRKIGVCTAKDGPAPCFIFGTVYRSGESFQSSC 120

Db 121 kyqctclldgavgcmlcsmdvrlpsdpfprvrklpgkceewvcdpkdqtvvvgpala 180
QY 121 KYQCTCLDgavgcmlcsmdvrlpsdpfprvrklpgkceewvcdpkdqtvvvgpala 180

Db 181 ayrledtfgpdpmtiranclvqtewsacsktcgmigistrvndnasrlekqslcmvr 240
QY 181 AYRL EDTFGPDPTMIRANCLVQTWEWSACSKTCGMIGISTRVNDNASCRLEKQSLCMVR 240

Db 241 pceadleenikkkgkcirtpkiskpikfelsgctsmktyrakfcgvcctdgrcctphrttt 300
QY 241 PCEADLEENIKKKGKCI RTPKISKPIKFELSGCTSMKTYRAKFCGVCCTDGRCTPHRTTT 300

Db 301 lpvefkcpdgevmkmmfiktacachncpgndndifeslyyrkmygdma 349
|||||

QY 301 LPVEFKCPDGEVVKNNMFIKTCACHYCPGNDNDFESLYYRKMVGMDMA 349

RESULT 3

ID R25366 standard; Protein; 348 AA.
 AC R25366;
 DT 18-JAN-1993 (first entry)
 DE Beta-IG-M2.
 KW Transforming growth factor beta; induced; CEF-10; v-src; chicken;
 KW embryo; fibroblasts; TGF-beta.
 OS Mus musculus.
 PN EP-495674-A.
 PD 22-JUL-1992.
 PF 17-JAN-1992; 300429.
 PR 18-JAN-1991; US-642991.
 PR 10-JAN-1992; US-816270.
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PI Brunner AM, Chinn J, Neubauer MG, Purchio AF;
 DR WFI; 92-243508/30.
 DR N-PSDB; Q26422.
 PT TGF-beta induced gene family - encodes proteins involved in
 PT growth and differentiation effects of TGF-beta-1
 PS Claim 3; Fig 2; 35pp; English.
 CC The protein sequence was deduced from the DNA sequence obtd. by
 CC screening a cDNA library made from AKR-2B mouse cells induced with
 CC TGF-beta1 and cyclohexamide with two probes from untreated AKR-2B
 CC mRNA and AKR-2B mRNA from cells treated with cyclohexamide and TGF-
 CC beta1. The proteins encoded by hybridising colonies (beta-IG-M1 and
 CC beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1.
 CC Beta-IG-M2 displays 50 percent homology to the CEF-10 protein
 CC induced by v-src in chicken embryo fibroblasts. Residues 52-59
 CC of beta-IG-M2 conform to the GCGCCXC motif reported in the
 CC amino half of insulin-like growth factor (IGF) binding proteins.
 CC The C-terminal Cys rich region of beta-IG-M1, -M2 and CEF-10 contain
 CC an amino acid sequence with strong homology to a motif found near the
 CC C-terminal of the malarial circumsporozoite (CS) protein, which is
 CC highly conserved among all species of malarial parasites sequenced
 CC to date (designated region II). This motif is also found in
 CC other proteins which have cell adhesive properties that mediate
 CC cell-cell and cell-extracellular matrix interactions, such as
 CC properdin, thrombospondin, and TRAP. The proteins encoded by
 CC TGF-beta induced genes are likely to be involved in mediation of
 CC the biological effects of TGF-beta relating to cell growth and
 CC differentiation. See also R25365.
 SQ Sequence 348 AA;

Query Match 92.9%; Score 2521; DB 5; Length 348;
 Best Local Similarity 91.1%; Pred. No. 4.65e-234;
 Matches 316; Conservative 21; Mismatches 9; Indels 1; Gaps 1;

Db 1 mlavagplslalv-llalctrpatgqdcasqccaaehpccagvsldvcgcrvc 59
 QY 1 MTAASMGPRVAFVLLALCSRPAGVQNCSPCRCPDEPAPCPAGVSLVLDGCGCRVC 60
 Db 60 akqlgelcterdpcdphkglfcdgspgpanrkigvctakdgapcvfgsvrsgesfqsc 119
 QY 61 AKQLGELCTERDPCDPHKGLFCDFGSPANRKIGVCTARDGAPCIFGTVYRSGESFQSSC 120
 Db 120 kvqctcldgavcvplcamdvrlbpdcpfrrvklpgkccceewvcdpdkrtavgpala 179
 |||||

QY 121 KYQCTCLDGAVGCMPLCSMDVRLPSDPDPFRVRKLPKGCCEEWCDQDTVGPALA 180
 Db 180 ayrlctfdgdpdmtmmrancelvqtewsacaktcmgmistrvtndntcfclqksrlcmvr 239
 QY 181 AYRLCTFDGDPDMTMRANCLVQTEWSACSKTCGMGISTRTVNDNASCLERKQSLCMVR 240
 Db 240 pceadleenikkckciirtpkiakpvkfelsgctavktyrakfcgvcgtgcrctphrttt 299
 QY 241 PCEADLEENIKKCKCIIRTPKISKPIKFELSGCTSMKTYRAKFCGVCCTDGRCTPHRTTT 300
 Db 300 lpvefkcpdgelmkkmmfiktcachnpcpgndndifeslyyrmymgdma 348
 QY 301 LPVEFKCPDGEVVKNNMFIKTCACHYCPGNDNDFESLYYRKMVGMDMA 349

RESULT 4

ID R31599 standard; Protein; 351 AA.
 AC R31599;
 DT 24-MAY-1993 (first entry)
 DE Chicken nov protein.
 KW avian nephroblastoma; avian myeloblastoma virus; IGF binding site;
 KW insulin-like growth factor; Wilm's tumour.
 OS Gallus domesticus.
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal peptide
 FT /note= "only hydrophobic region of protein"
 FT Binding site 56..63
 FT /label= IGF-binding site motif
 FT /note= "corresponds to GCGCCXC consensus"
 EN W09300430-A.
 PD 07-JAN-1993.
 PF 25-JUN-1992; F00589.
 PR 25-JUN-1991; FR-007807.
 PA (CNRS) CENT NAT RECH SCI.
 PI Martinerie C, Perbal B;
 DR WFI; 93-036377/04.
 DR N-PSDB; Q36031.
 PT Nucleotide sequences hybridising to regions of chicken nov gene -
 PT useful as probes for detecting complementary sequences to
 PT evaluate development and/or differentiation of tumours
 PS Claim 1; Fig 1; 67pp; French.
 CC This amino acid sequence was deduced from the nucleotide sequence
 CC of a chicken nov gene clone isolated from a gene bank prepared from
 CC chicken embryonic fibroblasts screened with a tumour-derived probe.
 CC The only hydrophobic region occurs within the putative signal
 CC peptide suggesting that the protein is secreted. The protein also
 CC contains the consensus motif of proteins which bind to insulin-like
 CC growth factors. It is known that the human IGFII gene is
 CC overexpressed in some Wilm's tumours and a similar deregulation of
 CC IGFII expression could be involved in nephroblastoma development.
 CC The deduced nov protein sequence contains 39 (non-clustered)
 CC cysteine residues.
 SQ Sequence 351 AA;

Query Match 53.8%; Score 1460; DB 6; Length 351;
 Best Local Similarity 54.2%; Pred. No. 4.62e-129;
 Matches 186; Conservative 69; Mismatches 80; Indels 8; Gaps 7;

Db 3 tgggggplpvllllllllrpcevsgraeacprcpggrcraep-prcapgvpavldgcgccl 61

```

QY 2 TAASMG-PVRVAVVLLALCSRPAGVQNCSGPC--RCDEPAPRCFAGVSLVLDGCGCCR 58
DQ 62 vcarqgescapllpodesaglycdrgpedggagimvlegdhcnvfdgmlyrngetfqp 121
QY 59 VCAKQLGELCTERDPCDPKHGLFCDFGSPANRKGIVCTAKDAGPCIFGTGYSRGSFQS 118
DQ 122 acyqctcrdggigclprcnlglppgdcprfkivpgeccckwcd-prdevllggf 180
QY 119 SKYQCTCLDGAAGVCMPLCSMDVRLSPDCPPFRVRKLPKGCCEWVCDPEKQTVVGP- 177
DQ 181 amaavqeatlgldvdaaancieqtwsacskcmgmfsftrvnrnqqcmvktqlrc 240
QY 178 ALAAAYLEDTFGDPFMIRANGLVQTEWSACSKTCGMGISTRVNDNASCRLEKQSRLC 237
DQ 241 mmrpeene-epedkkgkciqtckamkavrfeykncvavtqtkpkyogldngdrcctphn 299
QY 238 MVRPCEADLEENIKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGCTDGRCCTPHR 297
DQ 300 tktiqlvfcpgagkflkpmmlintevchngncpqnannaffqpl 342
QY 298 TTTLPVEFKCPDGEVKNMFMFKTCACHYNCP-GDNDIFESL 339

```

RESULT 5

```

ID R25565 standard; Protein; 379 AA.
AC R25565;
DE 18-JAN-1993 (first entry)
DE Beta-IG-M1.
KW Transforming growth factor beta; induced; CEF-10; v-src; chicken;
KW embryo; fibroblasts; TGF-beta.
OS Mus musculus.
PN EP-495674-A.
PD 22-JUL-1992.
PF 17-JAN-1991; US-642991.
PR 18-JAN-1991; US-816270.
PR 10-JAN-1992; US-816270.
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
PI Brunner AM, Chinn J, Neubauer MG, Purchio AF;
DR WPI; 92-243508/30.
DR N-PSDB; Q26421.
PT TGF-beta induced gene family - encodes proteins involved in
PT growth and differentiation effects of TGF-beta-1
PS Claim 2; Fig 1; 3pp; English.
CC The protein sequence was deduced from the DNA sequence obt. by
CC screening a cDNA library made from AKR-2B mouse cells induced with
CC TGF-beta1 and cyclohexamide with two probes from untreated AKR-2B
CC mRNA and AKR-2B mRNA from cells treated with cyclohexamide and TGF-
CC beta1. The proteins encoded by hybridising colonies (beta-IG-M1 and
CC beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1.
CC Beta-IG-M1 displays 80 percent homology to the CEF-10 protein
CC induced by v-src in chicken embryo fibroblasts and is identical
CC to the protein encoded by cyr61, an immediate early response gene
CC induced in quiescent BALB 3T3 cells by serum treatment. Residues
CC 49-56 of beta-IG-M1 conform to the GGCXMC motif reported in the
CC amino half of insulin-like growth factor (IGF) binding proteins.
CC The C-terminal Cys rich region of beta-IG-M1, -M2 and CEF-10 contain
CC an amino acid sequence with strong homology to a motif found near the
CC C-terminal of the malarial circumsporozoite (CS) protein, which is
CC highly conserved among all species of malarial parasites sequenced

```

to date (designated region II). This motif is also found in other proteins which have cell adhesive properties that mediate cell-cell and cell-extracellular matrix interactions, such as properdin, thrombospondin, and TRAP. The proteins encoded by TGF-beta induced genes are likely to be involved in mediation of the biological effects of TGF-beta relating to cell growth and cell differentiation. See also R25566.

SQ Sequence 379 AA;

```

Query Match 41.0%; Score 1113; DB 5; Length 379;
Best Local Similarity 48.3%; Pred. No. 5,50e-95;
Matches 180; Conservative 79; Mismatches 76; Indels 38; Gaps 20;

DQ 1 mssatftlava-vtllhl-trials-tcpaachcple-akcapagvlvrdgcgckvc 56
QY 1 MTAASMGFVRVAVVLLALCSRPAGVQNCSGPCRCDEPAPRCFAGVSLVLDGCGCCRVC 60
DQ 57 akqinedcsktqpcdhtkglecnfgaastalkgicraqeagrpceynstiyngesfqn 118
QY 61 AKQLGELCTERDPCDPKHGLFCDFGSPANRKGIVCTAK-DGAPCIFGTGYSRGSFQS 119
DQ 117 ckhctcidavgoipicpqlcpqlelplnlgcpnprlvkvsgccceewvcdedsikdalddgd 176
QY 120 KKYQCTCLDGAAGVCMPLCSMDVRLSPDCPPFRVRKLPKGCCEWVCDPE-P-KD----QT 173
DQ 177 dlglaseveltnneliaigkgsalkrlpvfqtprvlnplhahgkciavqttswaq 236
QY 174 -WVG--PA-----LA-----AY-R---LE--DTFGPDP-TM---IRAN---CLVQTWESA 208
DQ 237 cskcgtgiatrvtndnpeclrvketriceivpcgvpvyssllkkgkckskkspvrf 296
QY 209 CSKTCGMGISTRVNDNASCRLEKQSRLCMWRECEADLEENIKGKKCIRTPKISKPIKF 268
DQ 297 tyagcsavkkyrpkycgscvdcrcctplqtrvtkmfrcedgemfaknvmimgskscnyn 356
QY 269 ELSGCTSMKTYRAKFCGCTDGRCCTPHRTTTLPLVEFKCPDGEVKNMFMFKTCACHYN 328
DQ 357 cphneasf-rly 368
QY 329 CPGDNDI-FESLY 340

```

RESULT 6

```

ID R90919 standard; Protein; 375 AA.
AC R90919;
DE 25-JUN-1996 (first entry)
DE Connective tissue growth factor-2.
KW CTGF-2; connective tissue growth factor-2; secreted protein;
KW cartilaginous growth; skeletal; embryo; cell growth; morphogenesis;
KW insulin-like growth factor; fibroblast growth factor; Cry61.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..24
FT /label= signal_peptide
FT Protein 25..375
FT /label= mature_protein
FN W09601896-A.
PD 25-JAN-1996.
PF 12-JUL-1994; U07736.
PR 12-JUL-1994; W0-U07736.

```

PA (HUMA-) HUMAN GENOME SCI INC.
PI Adams MD, Li H;
DR WPI; 96-097626/10.
DR N-PSDB; T12653.
PT Connective tissue growth factor-2 and DNA encoding it - useful to
PT enhance the repair of connective and support tissue, and to enhance
PT wound healing
PS Claim 1; Fig 1A-C; 46pp; English.
CC Connective tissue growth factor-2 (CTGF-2) is encoded by a cDNA
CC (T12653) isolated from a human foetal lung cDNA library. The CTGF
CC polypeptides are structurally and functionally related to a family
CC of growth factors which include IGF (Insulin-like growth factor),
CC PDGF (platelet-derived growth factor), and FGF (fibroblast growth
CC factor). CTGF-2 exhibits 89 percent identity and 93 percent similarity
CC to Cry61. Cry61 is a growth factor-inducible immediate early gene
CC initially identified in serum-stimulated mouse fibroblasts. It encodes
CC a member of an emerging family of secreted proteins which are also a
CC group of cysteine-rich proteins. This group of GFs are important for
CC normal growth, differentiation, morphogenesis of the cartilaginous
CC skeleton of an embryo and cell growth.
SQ Sequence 375 AA;

Query Match 36.9%; Score 1002; DB 16; Length 375;
Best Local Similarity 47.2%; Pred. No. 3.88e-84;
Matches 175; Conservative 71; Mismatches 84; Indels 41; Gaps 21;
Db 6 vrelalvtllhl-trvqls-tcpadchcple-apkcapgvglvrdgcgcckvacakqlne 62
Qy 9 VR-VAFVV-LTALCSRPVAVGQNSGFCRCPDEPAPCPAGVSLVLDGCGCRCAKQLGE 66
Db 63 dorktpodhtkglecnfaastalkicraqsegrpceynsriyngesfqnckhqt 122
Qy 67 LCTERDPCDPKHLGFCDFGSPANRKGICTAK-DGAPCIFGGTVYRSGESFQSSCKYQCT 125
Db 123 cijwrrgacipclpelaipnlgcpnrlvkvgtgqceewvcddedskjmedqdgllgk 182
Qy 126 CLDGAVG-CMPLCSMDVRLVSPDCFPFRVRLPGRCCEEWVCDE----PK-DQT-VVGP 177
Db 183 glqfdaseveltrnnellavgkgsalkrlpvfmgprilypnlggkciqvttsweqcsk 242
Qy 178 AL---AA---Y-R---L-E---DT-----FGDPP-TM---IRA-NCLVQTTEWSACSK 211
Db 243 tegtgitrvtdnpeclrvketricvpcgppvyslkgkckstkkspepvrftya 302
Qy 212 TCGMGISTRVTDNNSCRLEQSLCVRPCEDADLEENIKKGGKCIPTPKISKPIKFLS 271
Db 303 gclsvkkyrpkycgcvgdgrctpqltrtkmrfpcedgetfaknmimgskcnyneph 362
Qy 272 GCTSMKTYRAKFCGVCITDGRCTPHRTTTLPEFKCPDGEVMKNNMFIKTCACHNCPG 331
Db 363 anea-afpyr 372
Qy 332 DNDIFESLYR 342
RESULT 7
ID R31608 standard; Protein; 205 AA.
AC R31608;
DT 24-MAY-1993 (first entry)
DE Homologous to chicken nov gene exon 3-4-encoded protein.

KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation; ss.
FN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PR (CNRS) CENT NAT RECH SCI.
PA Martinerie C, Perbal B;
PI WPI; 93-036377/04.
DR Nucleotide sequences hybridising to regions of chicken nov gene -
DR useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 21; Page 39; 67pp; French.

CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Fragment XXII (Q36044) is part of the 3rd and 4th exons of
CC the nov gene; nucleotide sequences which hybridise to Fragment XXII
CC under stringent conditions (i.e. 50% formamide, 5 x SSC) are claimed.
CC The claimed sequences preferably encode a protein with the sequence
CC XXIII (R31608) or an amino acid sequence 60% homologous to it.
SQ Sequence 205 AA;

Query Match 23.7%; Score 644; DB 6; Length 205;
Best Local Similarity 98.8%; Pred. No. 1.93e-49;
Matches 81; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 33 akdgapciftgtvyragesfgsackvgctcldgavcgmplcsmdvrlpsdpcpfrvkl 92
Qy 97 AKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSMDVRLPSDPCFPFRV 156
Db 93 pgkceewvcdepkdgtvlgpa 114
Qy 157 PGKCEEWVCDEPKDQTVVGP 178

RESULT 8
ID R31610 standard; Protein; 72 AA.
AC R31610;
DT 24-MAY-1993 (first entry)
DE Fragment XXVI with homology to chicken nov exon 3-encoded protein.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Homo sapiens.
FN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PR (CNRS) CENT NAT RECH SCI.
PA Martinerie C, Perbal B;
PI WPI; 93-036377/04.
DR Nucleotide sequences hybridising to regions of chicken nov gene -
DR useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 25; Page 40; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived

CC probe. Nucleotide sequences of the invention include those which
CC code for a protein having at least 60% homology with the protein nov
CC fragment XXV (R31609) deduced from the 3rd. exon of the chicken nov
CC gene. In particular, the sequences of the invention encode a
CC protein having the sequence of Fragment XXVI (R31610). (Sequence
CC XXVI is described as a nucleotide sequence in the claims but is
CC correctly described as an amino acid sequence in the disclosure).
SQ Sequence 72 AA;

Query Match 21.3%; Score 579; DB 6; Length 72;
Best Local Similarity 97.2%; Pred. No. 3.25e-43;
Matches 70; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 1 dgapcifggtvyrsgesfqscckycctclidgavgcmlcsmdvrlpsdpcfprrvklpg 60
Qy 99 DGAPCIFG GTVYRSGESFQSSCKYQCTCLIDGAVGCMPLCSMDVRLPSPDCFPFRRVKLPG 158
Db 61 mceewwvcdpr 72
Qy 159 KCCEEWVCDPEK 170

RESULT 9
ID R31602 standard; Protein; 84 AA.
AC R31602;
DT 24-MAY-1993 (first entry)
DE Chicken nov protein fragment IX.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Gallus domesticus.
PN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C. Perbal B;
DR WPI; 93-036377/04.
DR N-PSDB; Q36035.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 8; Page 31, 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence was isolated from a gene bank prepared from
CC chicken embryonic fibroblasts screened with a tumour-derived probe.
CC Fragment VIII (Q36035) is derived from the 3rd. exon of the nov gene;
CC nucleotide sequences which hybridise to Fragment VIII under stringent
CC conditions (i.e. 50% formamide, 5 x SCC) are claimed. The claimed
CC sequences preferably encode a protein with at least 70% homology to
CC amino acid sequence IX (R31602) which is encoded by Fragment VIII.
SQ Sequence 84 AA;

Query Match 15.8%; Score 430; DB 6; Length 84;
Best Local Similarity 57.7%; Pred. No. 4.30e-29;
Matches 45; Conservative 19; Mismatches 13; Indels 1; Gaps 1;

Db 3 egdncvfdgmlyrnetgfqpsckycqctordgicgclprcnlgllpgdpcfprrkivpg 62
Qy 99 DGAPCIFG GTVYRSGESFQSSCKYQCTCLIDGAVGCMPLCSMDVRLPSPDCFPFRRVKLPG 158

Db 63 eccekwvcd-prdevillg 79
Qy 159 KCCEEWVCDPEKDTWVG 176
RESULT 10
ID R31609 standard; Protein; 71 AA.
AC R31609;
DT 24-MAY-1993 (first entry)
DE Encoded by chicken nov gene exon 3 fragment.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Gallus domesticus.
PN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C. Perbal B;
DR WPI; 93-036377/04.
DR Nucleotide sequences hybridising to regions of chicken nov gene -
DR useful as probes for detecting complementary sequences to
DR evaluate development and/or differentiation of tumours
PS Claim 25; Page 40, 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Nucleotide sequences of the invention include those which
CC code for a protein having at least 60% homology with the protein
CC fragment XXV (R31609) deduced from the 3rd. exon of the chicken nov
CC gene. (Sequence XXV is described as a nucleotide sequence in the
CC claims but is correctly described as an amino acid sequence in the
CC disclosure).
SQ Sequence 71 AA;

Query Match 15.2%; Score 412; DB 6; Length 71;
Best Local Similarity 59.7%; Pred. No. 2.09e-27;
Matches 43; Conservative 16; Mismatches 12; Indels 1; Gaps 1;

Db 1 egdncvfdgmlyrnetgfqpsckycqctordgicgclprcnlgllpgdpcfprrkivpg 60
Qy 99 DGAPCIFG GTVYRSGESFQSSCKYQCTCLIDGAVGCMPLCSMDVRLPSPDCFPFRRVKLPG 158
Db 61 eccekwvcd-pr 71
Qy 159 KCCEEWVCDPEK 170

RESULT 11
ID R31603 standard; Protein; 84 AA.
AC R31603;
DT 24-MAY-1993 (first entry)
DE Polypeptide X homologous to chicken nov protein fragment.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Homo sapiens.
PN W09300430-A.
PD 07-JAN-1993.

Wed Sep 17 11:16:03 1997

Using for Mary Hale

PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C, Perbal B;
DR WPI; 93-036377/04.
DR N-PSDB; Q36036.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 10; Page 31; 67pp; French.
CC Nucleotide sequences which hybridise to Fragment VIII (Q36035) of
CC the 3rd. exon of chicken nov gene under stringent conditions (i.e.
CC 50% formamide, 5 x SCC) are claimed. The claimed sequences preferably
CC encode a protein with at least 70% homology to R31602 which is encoded
CC by Fragment VIII. Such nucleotide sequences are further
CC characterised in that they comprise at least part of a 700bp PstI
CC fragment derived from a recombinant clone isolated from a human
CC placenta DNA library and encode the amino acid sequence of
CC fragment X (R31603).
SQ Sequence 84 AA;

Query Match 14.9%; Score 405; DB 6; Length 84;
Best Local Similarity 53.7%; Pred. No. 9.43e-27;
Matches 44; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

Db 3 egdnvfdgvlvragekfapackfctdrcdgcvgvrcqldvllpepncpaprkvvepg 62
QY 99 DGAPCIFGTVYRSGESFQSSCKYQCTCLDGVAGCMPLCSMDVRLPSPDCFPFRVKLP 158

Db 63 eccekwicpdeedalgltla 84

QY 159 KCCEWVCDKFDQIVVGFALA 180

RESULT 12

ID R46078 standard; Protein; 124 AA.
AC R46078;
DT 19-OCT-1994 (first entry)
DE CYR61 like protein.
KW Human cDNA; library; enzyme; protein.
OS Homo sapiens.
PN W09403599-A.
PD 17-FEB-1994.
PF 04-AUG-1993; J01095.
PR 04-AUG-1992; JP-208077.
PR 13-NOV-1992; JP-327619.
PR 26-FEB-1993; JP-061431.
PA (SAGA) SAGAMI CHEM RES CENTRE.
PI Iwahori A, Kato S, Kato I, Kim N, Oh S, Sekine S;
DR WPI; 94-065688/08.
DR N-PSDB; Q57417.
PT cDNA of human origin and proteins coded by it - which may be
PT expressed by in vivo or in vitro translation using sense RNA or
PT antisense DNA corresponding to the cDNA.
PS Claim 1; Page 32-33; 167pp; Japanese.
CC mRNA expressed in human fibrosarcoma cell line HT-1080 was
CC isolated and used to construct a cDNA library using vector
CC pKA1. Clone HP00021 encoding CYR61-like protein
CC was isolated.
SQ Sequence 124 AA;

Wed Sep 17 11:16:03 1997

Using for Mary Hale

Query Match 14.6%; Score 395; DB 10; Length 124;
Best Local Similarity 58.4%; Pred. No. 8.10e-26;
Matches 45; Conservative 16; Mismatches 15; Indels 1; Gaps 1;
Db 6 gicragsegprceynariyngesfgpnckhctcldgavgcipicpgelelplngcnp 65
QY 93 GVTAK-DGAPCIFGTVYRSGESFQSSCKYQCTCLDGVAGCMPLCSMDVRLPSPDCFP 151
Db 66 rlkvvggqceewvdc 82
QY 152 RRVKLPKGCCEWVDC 168

RESULT 13

ID R31603 standard; Protein; 72 AA.
AC R31603;
DT 24-MAY-1993 (first entry)
DE Homologous to chicken nov exon 4-encoded protein.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringant hybridisation; CTGF.
PN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C, Perbal B;
DR WPI; 93-036377/04.
DR N-PSDB; Q36039.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 14; Page 34; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Fragment XIII (Q36038) is derived from the 4th. exon of the
CC nov gene; nucleotide sequences which hybridise to Fragment XIII
CC under stringent conditions (i.e. 50% formamide, 5 x SCC) are claimed.
CC The claimed sequences preferably encode a protein with at least 86%
CC homology to R31604 which is encoded by Fragment XIII. Pref. the
CC sequence of the invention comprises a sequence which encodes amino
CC acid sequence XV (R31603).
SQ Sequence 72 AA;

Query Match 11.6%; Score 314; DB 6; Length 72;
Best Local Similarity 58.3%; Pred. No. 2.49e-18;
Matches 42; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

Db 1 ayrepatlgvevdsavncieqtetwatsckscgmgtftrvtnnrqemlqtrlcmmvr 60
QY 181 AYRLETFGPDPTMIRANCLVQTTESACSKTCGNGISTRTVNDNASCRLEKQSLCVR 240

Db 61 pceqepetpdk 72

QY 241 PCEADLEENIKK 252

RESULT 14

ID R31604 standard; Protein; 70 AA.
AC R31604;
DT 24-MAY-1993 (first entry)
DE Chicken nov protein fragment XIV.
KW avian myeloblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Gallus domesticus.
PN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C. Perbal B;
DR WPI; 93-036377/04.
DR N-PSDB; Q36038.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 13; Page 34; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence was isolated from a gene bank prepared from
CC chicken embryonic fibroblasts screened with a tumour-derived probe.
CC Fragment XIII (Q36038) is derived from the 4th. exon of the nov gene;
CC nucleotide sequences which hybridise to Fragment XIII under stringent
CC conditions (i.e. 50% formamide, 5 x SCC) are claimed. The claimed
CC sequences preferably encode a protein with at least 86% homology to
CC amino acid sequence XIV (R31604) which is encoded by Fragment XIII.
SQ Sequence 70 AA;

Query Match 11.5%; Score 312; DB 6; Length 70;
Best Local Similarity 64.6%; Pred. No. 3.79e-18;
Matches 42; Conservative 9; Mismatches 14; Indels 0; Gaps 0;
Db 1 avrqeatlqdvdsasanciegttewasacskcmgfgstvtvtrnngqcmvkgtrclcmvr 60
QY 181 AYRLEDITGPDTMIRANCLVQTTEWSACSKTCGIGSTRVTNDNASCRLEKQSRCLCMVR 240
Db 61 pcene 65
QY 241 PCEAD 245

RESULT 15
ID R31601 standard; Protein; 75 AA.
AC R31601;
DT 24-MAY-1993 (first entry)
DE Chicken nov protein fragment V.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation.
OS Gallus domesticus.
PN W09300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C. Perbal B;
DR WPI; 93-036377/04.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to

PT evaluate development and/or differentiation of tumours
PS Claim 5; Page 28; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Nucleotide sequences which hybridise to Q36031 or specified
CC sub-fragments of it, under stringent conditions (i.e. 50% formamide,
CC 5 x SCC), are claimed. The claimed sequences preferably encode a
CC protein with amino acid sequence V (R31601).
SQ Sequence 75 AA;
Query Match 10.2%; Score 278; DB 6; Length 75;
Best Local Similarity 48.0%; Pred. No. 4.66e-15;
Matches 36; Conservative 14; Mismatches 22; Indels 3; Gaps 2;
Db 2 aatgcpqpqgpcrpatp-ptcagpgvrvldgcscclvcargcsdlepccdesagly 60
QY 24 AVGNQCSGPC--RCPDEPAPRCPCAGVSLVLDGGCCRCVCAKQKGELCTERDFCDPHKGLF 81
Db 61 cdrsadspsndtgict 75
QY 82 CDFGSPANRKIGVCT 96
Search completed: Wed Sep 17 09:36:20 1997
Job time : 72 secs.

maryh@stic

4380-1

NeWSprinter20

Wed Sep 17 11:17:00 1997

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5

MPARCH_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 17 09:33:28 1997; MaePar time 13.07 Seconds
Tabular output not generated.
771.467 Million cell updates/sec

Title: >US-08-167-628-2
Description: (1-349) from 5408040.ppe
Perfect Score: 2713
Sequence: 1 MTAASGMPVRVAFVILLALC.....PGDNDIFESLYRMYGDMA 349

Scoring table:
PAM 150
Gap 11

Searched: 91006 seqs, 28888923 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: .pir51
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
13:unann9 14:unann10 15:unann16:unrev

Statistics: Mean 42.720; Variance 82.202; scale 0.520

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description	Pred. No.
1	2713	100.0	349 13	A40551	connective tissue gr 0.00e+00
2	2521	92.9	348 14	A40578	beta IG-M2 protein p 0.00e+00
3	2514	92.7	348 14	A53228	fisp-12 protein prec 0.00e+00
4	1460	53.8	351 13	S20078	NOV protein - chicke 3.42e-284
5	1415	52.2	357 13	I38069	gene novH protein - 2.53e-274
6	1152	42.5	375 13	A41428	CEF-10 protein precu 7.43e-217
7	1113	41.0	379 14	A35669	gene CYR61 protein p 2.27e-208
8	133	4.9	254 14	JC1464	insulin-like growth 1.32e-06

9	133	4.9	254 14	I48599	insulin-like growth 1.32e-06
10	128	4.7	271 13	JC1484	insulin-like growth 8.02e-06
11	125	4.6	271 14	JC1484	insulin-like growth 8.02e-06
12	126	4.6	835 13	JP0076	nel protein - chicke 1.63e-05
13	123	4.5	258 13	B37252	insulin-like growth 4.71e-05
14	121	4.5	455 6	A43918	TGF-beta-related pro 9.49e-05
15	116	4.3	254 14	I48603	insulin-like growth 5.32e-04
16	117	4.3	277 13	I52825	gene MAC25 protein - 3.78e-04
17	117	4.3	282 13	S50031	prostaglandin-stimula 3.78e-04
18	117	4.3	1348 12	A43917	probable epidermal c 2.68e-04
19	118	4.3	1348 12	S27812	probable epidermal c 2.68e-04
20	115	4.2	272 6	A36082	insulin-like growth 7.48e-04
21	113	4.2	2813 3	VWHD	von Willebrand facto 1.47e-03
22	110	4.1	263 6	S23009	insulin-like growth 4.01e-03
23	111	4.1	271 14	I48604	insulin-like growth 2.88e-03
24	111	4.1	304 6	A33274	insulin-like growth 2.88e-03
25	111	4.1	305 14	I48601	insulin-like growth 2.88e-03
26	111	4.1	305 6	JN0508	insulin-like growth 2.88e-03
27	110	4.1	3020 13	A43932	mucin 2 precursor, i 4.01e-04
28	109	4.0	124 6	S25113	insulin-like growth 5.59e-03
29	109	4.0	272 14	I48600	insulin-like growth 5.59e-03
30	106	3.9	258 14	A45403	insulin-like growth 1.49e-02
31	105	3.9	272 13	A53748	insulin-like growth 2.07e-02
32	106	3.9	807 14	A38152	F-spondin - rat 1.49e-02
33	102	3.8	259 2	IOH01	insulin-like growth 5.40e-02
34	102	3.8	291 2	JN0064	insulin-like growth 5.40e-02
35	104	3.8	310 6	A60967	insulin-like growth 2.85e-02
36	103	3.8	328 2	A41927	insulin-like growth 3.93e-02
37	103	3.8	500 4	PLWL6	IL1 protein - human p 3.93e-02
38	102	3.8	798 7	A40526	integrin beta-7 chai 5.40e-02
39	104	3.8	3133 16	S70920	lectin - silkworm 2.85e-02
40	104	3.8	3133 16	S52093	circumsporozoite pro 7.42e-02
41	101	3.7	388 7	A39756	prepro osteogenic pr 7.42e-02
42	101	3.7	402 6	A45056	circumsporozoite pro 7.42e-02
43	101	3.7	405 7	S05428	circumsporozoite pro 7.42e-02
44	101	3.7	412 3	OZ20AF	circumsporozoite pro 7.42e-02
45	101	3.7	424 12	A54533	circumsporozoite pro 7.42e-02

ALIGNMENTS

RESULT	1
ENTRY	A40551
TITLE	connective tissue growth factor - human
ORGANISM	#formal name Homo sapiens #common name man
DATE	17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 25-Aug-1995
ACCESSIONS	A40551; S44205
REFERENCE	A40551
#authors	Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotendorst, G.R.
#journal	J. Cell Biol. (1991) 114:1285-1294
#title	Connective tissue growth factor: a cysteine-rich mitogen secreted by human vascular endothelial cells is related to the SRC-induced immediate early gene product CEF-10.
#cross-references	W0191373462
#accession	A40551
#molecule type	mRNA
##residues	1-349
#label	BRA
REFERENCE	S44205
#authors	Oemar, B.S.; Werner, A.; Yang, Z.; Garnier, J.M.; Gentz, R.;

#submission Luescher, T.F.
#description submitted to the EMBL Data Library, April 1994
#accession Differential cloning and expression of human connective
#status tissue growth factor.
#molecule_type preliminary
#residues 544205
#cross-references mRNA
#length 349 #molecular-weight 38069 #checksum 8930

Query Match 100.0%; Score 2713; DB 13; Length 349;
Best Local Similarity 100.0%; Pred. No. 0.00e+00; Indels 0; Gaps 0;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mtaasgprvrvafvllalcalcarpavggncgpcrcpdepaprcpagvslvldgcgcrcv 60
Qy 1 MTAASGPRVRVAFVLLALCSRPVAVGQNCSPCRCPDEPAPRCAGVSLVLDGCGCRCV 60
Db 61 akqlgelcterdpchkhgkglfcdgspgpanrkigvctakdgapcfiggtvyrsgesfgasc 120
Qy 61 AKQLGELCTERDPCDHPKHGKGLFCDFGSPANRKIGVCTAKDGPACFIGGTVYRSGESFGASC 120
Db 121 kyqctcldgavgcplcsmdvrlpsdpcpfprvrklpgkccceewvcdpdkdtvvgpala 180
Qy 121 KYQCTCLDGA VGCPLCSMDVRLPSDPCPFPRVRKLPGKCCCEEWVCDPDKDTVVGPALA 180
Db 181 avrledtfgdpdmirancvlqtetwascsktcgmgiatrvtdnnaacrlkgsrlcmvr 240
Qy 181 AYRLEDTFGDPDMIRANCLVQTETWASCSKTCGMGIATRVTDNNAACRLKGSRLCMVR 240
Db 241 pceadleenikkkgkciirtpkiakpvkfelsagtsmktyrakfgvctdgrctphrttt 300
Qy 241 PCEADLEENIKKKGKCIIRTPKISKPIKELSGTSMKTYRAKFGVCTDGRCTPHRTTT 300
Db 301 lpvefkcpdgvemkmmfiktacachyncpgndndifeslyyrkmygdma 349
Qy 301 LPVEFKCPDGEVMKMMFIKTCACHYNCPGNDNDFESLYYRKMYGDMA 349

RESULT 2 A40578 #type complete
ENTRY beta IG-M2 protein precursor - mouse
TITLE #formal_name Mus musculus #common name house mouse
ORGANISM 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change
DATE 30-Sep-1993
ACCESSIONS A40578
REFERENCE A40578
#authors Brunner, A.; Chinn, J.; Neuberger, M.; Purchio, A.F.
#journal DNA Cell Biol. (1991) 10:293-300
#title Identification of a gene family regulated by transforming
#cross-references MUID:91229699
#accession A40578
#status preliminary
#molecule_type mRNA
#residues 1-348 #label BRU
#length 348 #molecular-weight 37794 #checksum 9188

Query Match 92.9%; Score 2521; DB 14; Length 348;

Best Local Similarity 91.1%; Pred. No. 0.00e+00;
Matches 318; Conservative 21; Mismatches 9; Indels 1; Gaps 1;

Db 1 mlaasvagiialv-llaictrpatgdcacqcaaeahpcpagvalvidgcgcrcv 59
Qy 1 MTAASGPRVRVAFVLLALCSRPVAVGQNCSPCRCPDEPAPRCAGVSLVLDGCGCRCV 60
Db 60 akqlgelcterdpchkhgkglfcdgspgpanrkigvctakdgapcfiggtvyrsgesfgasc 119
Qy 61 AKQLGELCTERDPCDHPKHGKGLFCDFGSPANRKIGVCTAKDGPACFIGGTVYRSGESFGASC 120
Db 120 kyqctcldgavgcplcsmdvrlpsdpcpfprvrklpgkccceewvcdpdkdtvvgpala 179
Qy 121 KYQCTCLDGA VGCPLCSMDVRLPSDPCPFPRVRKLPGKCCCEEWVCDPDKDTVVGPALA 180
Db 180 avrledtfgdpdmirancvlqtetwascsktcgmgiatrvtdnndtforlekqslcmvr 239
Qy 181 AYRLEDTFGDPDMIRANCLVQTETWASCSKTCGMGIATRVTDNNAACRLKGSRLCMVR 240
Db 240 pceadleenikkkgkciirtpkiakpvkfelsagtsmktyrakfgvctdgrctphrttt 299
Qy 241 PCEADLEENIKKKGKCIIRTPKISKPIKELSGTSMKTYRAKFGVCTDGRCTPHRTTT 300
Db 300 lpvefkcpdgvemkmmfiktacachyncpgndndifeslyyrkmygdma 348
Qy 301 LPVEFKCPDGEVMKMMFIKTCACHYNCPGNDNDFESLYYRKMYGDMA 349

RESULT 3 A53228 #type complete
ENTRY flisp-12 protein precursor - mouse
TITLE #formal_name Mus musculus #common name house mouse
ORGANISM 19-May-1994 #sequence_revision 19-May-1994 #text_change
DATE 19-May-1994
ACCESSIONS A53228
REFERENCE A53228
#authors Rybeck, R.P.; Macdonald-Bravo, H.; Mattei, M.G.; Bravo, R.
#journal Cell Growth Differ. (1991) 2:225-233
#title Structure, mapping, and expression of flisp-12, a growth
factor-inducible gene encoding a secreted cysteine-rich
protein.
#accession A53228
#status preliminary
#molecule_type DNA
#residues 1-348 #label RYS
#cross-references GB:M70641

GENETICS
#gene flisp-12
SUMMARY #length 348 #molecular-weight 37793 #checksum 9470
Query Match 92.7%; Score 2514; DB 14; Length 348;
Best Local Similarity 90.8%; Pred. No. 0.00e+00;
Matches 317; Conservative 21; Mismatches 10; Indels 1; Gaps 1;

Db 1 mlaasvagiialv-llaictrpatgdcacqcaaeahpcpagvalvidgcgcrcv 59
Qy 1 MTAASGPRVRVAFVLLALCSRPVAVGQNCSPCRCPDEPAPRCAGVSLVLDGCGCRCV 60
Db 60 akqlgelcterdpchkhgkglfcdgspgpanrkigvctakdgapcfiggtvyrsgesfgasc 119
Qy 61 AKQLGELCTERDPCDHPKHGKGLFCDFGSPANRKIGVCTAKDGPACFIGGTVYRSGESFGASC 120

QY 61 AKQLGELTERDPCDPHGKLFCDGSPANRKGIVCTAKDGNAPCIFGCTVYRSGESFQSSC 120

Db 120 kyqctcldgavgcplcsmdvrlpsdcpfprrvklpgkckewwvdepkdrtaavgpala 179

QY 121 KYQCTCLDGAVGCMPLCSMDVRLPSDPCFPFRRVKLPKGCCEEWVCDPKDQTVWGPALA 180

Db 180 ayrltedtfgpdpmtmrancivqtteasacsktcmgmistrvntndntfcirlekgserlcmvr 239

QY 181 AYRLTEDTFGDPDMTMRANCLVQTTEWSACSKTCGMGISTRTVNDNASCRLEKQSRCLMVR 240

Db 240 pceadleenikgkciirtkpkakpvkfelsgetavtktyrakfcgvtcdgrectphrttt 299

QY 241 PCEADLEENIKGKCIIRTPKISPKIFELSGCTSMKTYRAKFCGVTCDGRCCTPHRTTT 300

Db 300 lpvefkcpdgelimkmmfiktacchyncpogndndifesyryrmygdma 348

QY 301 LPVEFKCPDGEVMKKNMFIKTCACHYNCPGDNDIFESLYRMYGDMA 349

RESULT 4

ENTRY S20078 #type complete

TITLE NOV protein - chicken

ORGANISM #formal name Gallus gallus #common name chicken

DATE 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995

ACCESSIONS S20078

REFERENCE Joliot, V.; Martinerie, C.; Dambrine, G.; Plassiart, G.;

#authors Brisac, M.; Crochet, J.; Perbal, B.

#journal Mol. Cell. Biol. (1992) 12:10-21

#title Proviral rearrangements and overexpression of a new cellular gene (nov) in myeloblastosis-associated virus type 1-induced nephroblastomas.

#cross-references MUID:92107157

#accession S20078

GENETICS

#status preliminary

#molecule_type mRNA

#residues 1-351 #label JOL

#cross-references EMBL:X59284

#gene NOV

SUMMARY

#length 351 #molecular-weight 38268 #checksum 4069

Query Match 53.8%; Score 1460; DB 13; Length 351;

Best Local Similarity 54.2%; Pred. No. 3.42e-284;

Matches 186; Conservative 69; Mismatches 80; Indels 8; Gaps 7;

Db 3 tgggggplvlllllllllpcvsgreaacprpcggrcpaep-prcagpvpavldgagcl 61

QY 2 TAASMG-PVRVAFVLLALCSRRPAGVQNCSPC--RCPDEPAPRCFAGVSLVDGGCCR 58

Db 62 vcarqgescapllpcdesaglycdrpdegaggagicmvlcdncvfdgmlyrnetfcp 121

QY 59 VCAGKQELGELTERDPCDPHGKLFCDGSPANRKGIVCTAKDGNAPCIFGGTVYRSGESFQ 118

Db 122 ackyqctcdgigclprcnlglilpgdpdpfprkievpgcecekwwcd-prdevllgff 180

QY 119 SKYQCTCLDGAVGCMPLCSMDVRLPSDPCFPFRRVKLPKGCCEEWVCDPKDQTVWGP- 177

Db 181 amaaYrqeatlgidvdsasancieqtteasacskacgmgtfstrvtrnngqcmvktqrlc 240

QY 178 ALAAYLEDTFGDPDMTMRANCLVQTTEWSACSKTCGMGISTRTVNDNASCRLEKQSRCLC 237

Db 241 mmrppcene-epedkkgkciqtktkamkavrfeykncvavqtkyprycglndgrectphn 299

QY 238 MVRPCEADLEENIKGKCIIRTPKISPKIFELSGCTSMKTYRAKFCGVTCDGRCCTPHR 297

Db 300 tktiqvfcrcpqqgkflkkpmlintevchngncpqnannaffqpl 342

QY 298 TTTLPVEFKCPDGEVMKKNMFIKTCACHYNCP-GDNDIFESL 339

RESULT 5

ENTRY I38069 #type complete

TITLE gene novH protein - human

ORGANISM #formal name Homo sapiens #common name man

DATE 17-May-1996 #sequence_revision 17-May-1996 #text_change 17-May-1996

ACCESSIONS I38069

REFERENCE Martinerie, C.; Huff, V.; Joubert, I.; Badzioch, M.;

#authors Saunders, G.; Strong, I.; Perbal, B.

#journal Oncogene (1994) 9:2729-2732

#title Structural analysis of the human nov proto-oncogene and expression in Wilms tumor.

#cross-references MUID:94336229

#accession I38069

GENETICS

#status preliminary; translated from GB/EMBL/DBJ

#molecule_type DNA

#residues 1-357 #label RES

#cross-references EMBL:X78351; NID:g587422; CDS_PID:g825696

GENETICS

#introns 28/3; 104/1; 188/1; 259/3

#note gene name novH

SUMMARY

#length 357 #molecular-weight 39162 #checksum 1640

Query Match 52.2%; Score 1415; DB 13; Length 357;

Best Local Similarity 51.8%; Pred. No. 2.53e-274;

Matches 177; Conservative 68; Mismatches 91; Indels 6; Gaps 5;

Db 18 ltfll-llhllggvaatgcpccpgrcnatp-ptcagpvrvldgscclvcargesc 75

QY 11 VAFVLLALCSRRPAGVQNCSPC--RCPDEPAPRCFAGVSLVDGGCCRCVAKQELGCLC 68

Db 76 sdlepdeasaglycdrsadpsnhtgtictavegdnvfdgviyrsgkfkpckfctcd 135

QY 69 TERDPCDPHGKLFCDGSPANRKGIVCTAKDGNAPCIFGGTVYRSGESFQSSCKYQCTCLD 128

Db 136 gqigcvproqldvllpepncpaprkvvepgcecekwicgpddeedalgltlaayrpeatl 195

QY 129 GAVGCMPLCSMDVRLPSDPCFPFRRVKLPKGCCEEWVCDPKDQTVWGPALAAVRLDFT 188

Db 196 gvevdsasvncieqtteasacskacgmgtfstrvtrnngqcmvktqrlcvmrpeceqeq 255

QY 189 GPDMTMRANCLVQTTEWSACSKTCGMGISTRTVNDNASCRLEKQSRCLMVRPCEADLEE 248

Db 256 ptdkkgkclrtkkskalkhlgfctalshtykrfcgvcsgdrectphntktqaeqfc 315

QY 249 NI-KKGGKCIIRTPKISPKIFELSGCTSMKTYRAKFCGVTCDGRCCTPHRTTLPVEFKC 307

Db	316	spgqivkvpvmvigtctchtnopknearfglelelkttrgkm	357
Qy	308	PDGEVMKKNNMFIKTKACHYNGFGNDIF-ESLYYRMYGDM	348
RESULT	6		
ENTRY	A14128	#type complete	
TITLE	CSF-10 protein precursor - chicken		
ORGANISM	#formal name Gallus gallus #common name chicken		
DATE	03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 05-May-1993		
ACCESSIONS	A14128		
REFERENCE	A14128		
#authors	Simmons, D.L.; Levy, D.B.; Yannoni, Y.; Erikson, R.L.		
#journal	Proc. Natl. Acad. Sci. U.S.A. (1989) 86:1178-1182		
#title	Identification of a phorbol ester-repressible v-src-inducible gene.		
#cross-references	MDID:89145206		
#accession	A14128		
#status	preliminary		
#molecule type	mRNA		
#residues	1-375 #label SIM		
#cross-references	GB:J04496		
SUMMARY	#length 375 #molecular-weight 40651 #checksum 1417		
Query Match	42.5%;	Score 1152; DB 13; Length 375;	
Best Local Similarity	48.5%;	Pred. No. 7.43e-217;	
Matches	176; Conservative	76; Mismatches 77; Indels 34; Gaps 19;	
Db	7	rpalaal-lclarialgpcpavcqpaa-aapcagvglvpgccckvcaqlnecd	64
Qy	10	RVAFVLLALC-SRPAVGNCSPGCRCPDEPAPCPAGVSLVLDGCCRCRCAKQLELC	68
Db	65	artqpedtkglecnfgaapaatngicraqegrpceynskiyngesfapnckhqtcti	124
Qy	69	TERDPCDPHKGLFCDFGSPANKRIGVCTAK-DGAPCIFGTVYRSGESFQSSCKQCTCL	127
Db	125	dgavgcipicpgeislplngcpsprlvkpgqceewvcdeskdaleegffskefgid	184
Qy	128	DGAVGCMPLCSMDVRLPSPDCFPRARVLPGKCEEWVCDPEKD--QTVVVG--P--AL-	179
Db	185	asegeltrnnelailavkggikmlpvgsepgsrafenpckivqttswsgcsktcgigt	244
Qy	180	AA---Y-R---I-E-----TFGDDP-TM-I-RANCLVQTTEWSACSKTCGMGIST	219
Db	245	rvtndnpdckllketricvevrcpgcpsyaslkkgkctktkpsprftyagcsavkky	304
Qy	220	RVTNDNASCRLEKQSRCLMVRPCEADLEENIKGKKCIRTPIKPIKFELSGCTSMKTY	279
Db	305	rpkygagcvdgcrcctpqqrtrtkifrcddgetftkvmmiqscronynophaneay-p-	362
Qy	280	RAKFCGCVTDGRCCTPHRTTLTLPVEFKPDGEVMKKNNMFIKTKACHYNGFGNDIFESL	339
Db	363	ivr	365
Qy	340	YR	342
RESULT	7		
ENTRY	A35669	#type complete	

```

TITLE
ORGANISM      gene CYR61 protein precursor - mouse
DATE          #format name Mus musculus #common name house mouse
              28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change
              18-Oct-1996
ACCESSIONS    A35669; I48319
REFERENCE     O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.
#journal      Mol. Cell. Biol. (1990) 10:3569-3577
#authors      Expression of cyr61, a growth factor-inducible
#title        immediate-early gene.
#cross-references MUID:90287146
#accession    A35669
##status      preliminary
##molecule_type mRNA
##residues    1-379 #label OAB
##cross-references GB:M32490
#note         the authors translated the codon GAT for residue 337
               Glu
REFERENCE     I48319
#authors      Latinkic, B.V.; O'Brien, T.P.; Lau, L.F.
#journal      Nucleic Acids Res. (1991) 19:3261-3267
#title        Promoter function and structure of the growth
               factor-inducible immediate early gene cyr61.
#cross-references MUID:91288203
#accession    I48319
##status      translated from GB/EMBL/DBJ
##molecule_type DNA
##residues    1-379 #label RES
##cross-references EMBL:X56790; NID:g50632; CDS_PID:g50633
GENETICS
#gene         CYR61
#introns      21/3; 93/1; 208/1; 279/3
SUMMARY       #length 379 #molecular-weight 41709 #checksum 3726
               Query Match 41.0%; Score 1113; DB 14; Length 379;
               Best Local Similarity 48.3%; Pred. No. 2,27e-208;
               Matches 180; Conservative 79; Mismatches 76; Indels 38; Gaps 20;
Db           1 masstfrtlava-vllhl-trlale-tcpaaachocple-apkcapgvglvrdgocckvc 56
QY           1 MTAAAGMPVRVAVFLLALCSRAVGQNGSGFCRCPEPAPPCGPAVSILVDGGCGRVC 60
Db           57 akqlnedosktqpcdhrtkglecnfgasatalkgicrasgprceynsrlyngesfqpn 116
QY           61 AKQLGELCTENDPCDPHKGLFLCDFGPANRK1GVCTAK--DGAPCIFGTGYRGSSEFSQS 119
Db           117 khqtctidagvgpicpgelalpnlgcnpnrlvkvgggcseewcdedsikdsiddq 176
QY           120 CKYQCTCLDAGVGCMPLCSMDVRLPSDFCFPRVKLGKCCCEWVCD-E-P-KD-----QT 173
Db           177 dlldlgdaeeveltrnnleiaigkssalrklypfgtgeprvlfnplahgqkcivqttswaq 236
QY           174 -VVG---PA---LA-----AY-R---LE---DTFGDP--TM---IRAN---CLVQTTEWSA 208
Db           237 cskcgdgistrvtndndpecrikvetricvrcpggpyvalkkgtcktkkspcvrf 296
QY           209 CSKTGGMGISITRVNDNASCRLEBKSLRWCPREADLEENIKKGKKICITPKISPIKF 268
Db           297 tyagsavkvkyprkyegscvdgrccptqlgtrvtvmkrfcedgemfsknvmigschcnyn 356

```

OV 269 ELSGCTSMKTYRAFCGVC TDGRCC TPRHTT TLPEFFKCPDGEVMKKNMMFIKTCACHYN 328

Db 357 cphpneasf-rly 368
|| |:
Qv 329 cpgdndi-fesly 340

RESULT	8
ENTRY	JC1464
TITLE	insulin-like growth factor-binding protein 4 precursor - rat
ORGANISM	#formal name Rattus norvegicus #common name Norway rat
DATE	30-Sep-1993 #sequence revision 30-Sep-1993 #text change 01-Dec-1995
ACCESSIONS	JC1464; A37252; E40403; A61120; B33570; A6118
REFERENCE	JC1464
#authors	Gao, L.; Ling, N.; Shimasaki, S.
#journal	Biochem. Biophys. Res. Commun. (1993) 190:1053-1059
#title	Structure of the insulin-like growth factor Binding protein-4 gene.

```
#accession JCI1464
#molecule_type DNA
#residues 1-254 ##label GAO
##cross-references GB:L08276
```

#authors	#journal	#title	#citations	#references
Shimasaki, S.; Uchiyama, F.; Shimonaka, M.; Ling, N. M.	Mol. Endocrinol. (1990) 4:1451-1458	Molecular cloning of the cDNAs encoding a novel insulin growth factor-binding protein from rat and human.	1	1

#cross-references W01D:91133415
#accession A37252
#molecule_type mRNA
#residues 1-254 ##label SHI
REFERENCE A04003
Shimaek, S.; Shimonaka, M.; Zhang, H.P.; Ling, N.
J. Biol. Chem. (1991) 266:10646-10653
#journal
#title Identification of five different insulin-like growth factor
binding proteins (IGFBPs) from adult rat serum and
molecular cloning of a novel IGFBP-5 in rat and human.
#cross-references W01D:91244847

```

#cross-references E0403
#accession E0403
##molecule_type protein
##residues 22-26,'X',28-28,'X',31-37,'X',39-40 ##label SH2

```

REFERENCE	#authors	#journal	#title
22-29, A, 26-27, A, 31-31, A, 33-40, 44-46, 48, 49, 51-52, 54-55, 57-58, 60-61, 63-64, 66-67, 69-70, 72-73, 75-76, 78-79, 81-82, 84-85, 87-88, 90-91, 93-94, 96-97, 99-100, 102-103, 105-106, 108-109, 111-112, 114-115, 117-118, 120-121, 123-124, 126-127, 129-130, 132-133, 135-136, 138-139, 141-142, 144-145, 147-148, 150-151, 153-154, 156-157, 159-160, 162-163, 165-166, 168-169, 171-172, 174-175, 177-178, 180-181, 183-184, 186-187, 189-190, 192-193, 195-196, 198-199, 201-202, 204-205, 207-208, 210-211, 213-214, 216-217, 219-220, 222-223, 225-226, 228-229, 231-232, 234-235, 237-238, 240-241, 243-244, 246-247, 249-250, 252-253, 255-256, 258-259, 261-262, 264-265, 267-268, 270-271, 273-274, 276-277, 279-280, 282-283, 285-286, 288-289, 291-292, 294-295, 297-298, 300-301, 303-304, 306-307, 309-310, 312-313, 315-316, 318-319, 321-322, 324-325, 327-328, 330-331, 333-334, 336-337, 339-340, 342-343, 345-346, 348-349, 351-352, 354-355, 357-358, 360-361, 363-364, 366-367, 369-370, 372-373, 375-376, 378-379, 381-382, 384-385, 387-388, 390-391, 393-394, 396-397, 399-400, 402-403, 405-406, 408-409, 411-412, 414-415, 417-418, 420-421, 423-424, 426-427, 429-430, 432-433, 435-436, 438-439, 441-442, 444-445, 447-448, 450-451, 453-454, 456-457, 459-460, 462-463, 465-466, 468-469, 471-472, 474-475, 477-478, 480-481, 483-484, 486-487, 489-490, 492-493, 495-496, 498-499, 501-502, 504-505, 507-508, 510-511, 513-514, 516-517, 519-520, 522-523, 525-526, 528-529, 531-532, 534-535, 537-538, 540-541, 543-544, 546-547, 549-550, 552-553, 555-556, 558-559, 561-562, 564-565, 567-568, 570-571, 573-574, 576-577, 579-580, 582-583, 585-586, 588-589, 591-592, 594-595, 597-598, 600-601, 603-604, 606-607, 609-610, 612-613, 615-616, 618-619, 621-622, 624-625, 627-628, 630-631, 633-634, 636-637, 639-640, 642-643, 645-646, 648-649, 651-652, 654-655, 657-658, 660-661, 663-664, 666-667, 669-670, 672-673, 675-676, 678-679, 681-682, 684-685, 687-688, 690-691, 693-694, 696-697, 699-700, 702-703, 705-706, 708-709, 711-712, 714-715, 717-718, 720-721, 723-724, 726-727, 729-730, 732-733, 735-736, 738-739, 741-742, 744-745, 747-748, 750-751, 753-754, 756-757, 759-760, 762-763, 765-766, 768-769, 771-772, 774-775, 777-778, 780-781, 783-784, 786-787, 789-790, 792-793, 795-796, 798-799, 801-802, 804-805, 807-808, 810-811, 813-814, 816-817, 819-820, 822-823, 825-826, 828-829, 831-832, 834-835, 837-838, 840-841, 843-844, 846-847, 849-850, 852-853, 855-856, 858-859, 861-862, 864-865, 867-868, 870-871, 873-874, 876-877, 879-880, 882-883, 885-886, 888-889, 891-892, 894-895, 897-898, 900-901, 903-904, 906-907, 909-910, 912-913, 915-916, 918-919, 921-922, 924-925, 927-928, 930-931, 933-934, 936-937, 939-940, 942-943, 945-946, 948-949, 951-952, 954-955, 957-958, 960-961, 963-964, 966-967, 969-970, 972-973, 975-976, 978-979, 981-982, 984-985, 987-988, 990-991, 993-994, 996-997, 999-1000, 1002-1003, 1005-1006, 1008-1009, 1011-1012, 1014-1015, 1017-1018, 1020-1021, 1023-1024, 1026-1027, 1029-1030, 1032-1033, 1035-1036, 1038-1039, 1041-1042, 1044-1045, 1047-1048, 1050-1051, 1053-1054, 1056-1057, 1059-1060, 1062-1063, 1065-1066, 1068-1069, 1071-1072, 1074-1075, 1077-1078, 1080-1081, 1083-1084, 1086-1087, 1089-1090, 1092-1093, 1095-1096, 1098-1099, 1101-1102, 1104-1105, 1107-1108, 1110-1111, 1113-1114, 1116-1117, 1119-1120, 1122-1123, 1125-1126, 1128-1129, 1131-1132, 1134-1135, 1137-1138, 1140-1141, 1143-1144, 1146-1147, 1149-1150, 1152-1153, 1155-1156, 1158-1159, 1161-1162, 1164-1165, 1167-1168, 1170-1171, 1173-1174, 1176-1177, 1179-1180, 1182-1183, 1185-1186, 1188-1189, 1191-1192, 1194-1195, 1197-1198, 1199-1200, 1202-1203, 1205-1206, 1208-1209, 1211-1212, 1214-1215, 1217-1218, 1220-1221, 1223-1224, 1226-1227, 1229-1230, 1232-1233, 1235-1236, 1238-1239, 1241-1242, 1244-1245, 1247-1248, 1250-1251, 1253-1254, 1256-1257, 1259-1260, 1262-1263, 1265-1266, 1268-1269, 1271-1272, 1274-1275, 1277-1278, 1280-1281, 1283-1284, 1286-1287, 1289-1290, 1292-1293, 1295-1296, 1298-1299, 1301-1302, 1304-1305, 1307-1308, 1310-1311, 1313-1314, 1316-1317, 1319			

#accession	#molecule type protein	#residues	#label	#note
A61120	'X', 23-26, 'X', 28-29, 'X', 31-37, 'X', 39-43, 'X', 45, 'X', 47-48		both glycosylated and nonglycosylated forms of this protein were observed	

REFERENCE

A33570

Shimonaka, M.; Schroeder, R.; Shimasaki, S.; Ling, N. Biochem. Biophys. Res. Commun. (1989) 165:189-195

#authors

#journal

#title

Identification of a novel binding protein for insulin-

growth factors in adult rat serum.
#cross-references MUID:90073708
#accession B33570
##molecule_type protein
##residues 22-26,'X',28-29,'X',31-37,'X',39-43,'X',45-52,'X',54,
'XX',57-58,'X',60-61 ##label SH3
REFERENCE A6118
#authors Cheung, P.T.; Smith, E.P.; Shimaaki, S.; Ling, N.;
Chernausek, S.D.
#journal Endocrinology (1991) 129:1006-1015
#title Characterization of an insulin-like growth factor binding
protein (IGFBP-4) produced by the B104 rat neuronal cell
line: chemical and biological properties and differential
synthesis by sublines.

REFERENCE
#authors

#journal
#title
Endocrinology (1991) 129:1006-1015
Characterization of an insulin-like growth factor binding protein (IGFBP-4) produced by the B104 rat neuronal cell line: chemical and biological properties and differential synthesis by sublines.

```
#accession A61118
##molecule_type protein
##residues 22
```

117/1; 165/3; 210/3
#introns
CLASSIFICATION
#superfamily thyroglobulin type I repeat hom

KEYWORDS	FEATURE	#domain signal sequence	#status experimental	#label ST
glycoprotein	1-21			
	22-254			

```

170-245
#status experimental #label MAT\
#domain thyroglobulin type I repeat homology #label
      THY1\

```

```

125 #binding_site carbohydrate (Asn) (covalent) #status
SUMMARY #length 254 #molecular-weight 27745 #checksum 4589

```

Query Match 4.9%; Score 133; DB 14; Length 254;
Best Local Similarity 39.7%; Pred. No. 1.32e-06;

Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6

Db 1 mlpfglvaalllaagprpslgde-ahhpcpcseeklarcrppvvgceelvrepaggccatc 59

Qy 6 MGPRVAFVVLALCSRPAVGQNCSGPCR-CPDEFAPRC-PA-GVS-LVID-GGCCRCVC 60

Db 60 aglgmpc 67
| | | | |
Qy 61 AKQLGELC 68

RESULT 9

ENTRY	#type complete
TITLE	insulin-like growth factor binding protein 4 - mouse
ORGANISM	#formal_name Mus musculus #common_name house mouse

DATE	02-Jul-1996	#sequence	revision	02-Jul-1996	#text_change
			18-Oct-1996		
ACCESSIONS		I48599;	S38668		
.....					

REFERENCE

I48599

Authors Bethel, C.R.; Vitullo, J.C.; Miller, R.E.; Aron, D.C.

Journal Biochem. Mol. Biol. Int. (1994) 34:385-392

[illegible]

```
#cross-references MUID:95152444
#accession I48599
##status preliminary; translated from GB/EMBL/DBJ
```



```

##molecule_type DNA
##residues 1-254 ##label RES
##cross-references EMBL:X76066; NID:g416032; CDS PID:g416033
CLASSIFICATION #superfamily thyroglobulin type I repeat homology
FEATURE
170-245 #domain thyroglobulin type I repeat homology #label THY1
SUMMARY
#length 254 #molecular-weight 27807 #checksum 3509
Query Match 4.9%; Score 133; DB 14; Length 254;
Best Local Similarity 39.7%; Pred. No. 1.32e-06;
Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6;
Db 1 mlpfglvaallaaqprslgde-aihhpcseeklarcrppvgceelvrepqgcgcac 59
Qy 6 MGPVRVAFVWLLALCSRPVAVGQNSGFC-CPDEPAPRC-PA-GVS-LVLD-GCGCCRCVC 60
Db 60 alqgmcp 67
Qy 61 AKQLGELC 68

RESULT 10
ENTRY JC4584 #type complete
TITLE insulin-like growth factor binding protein-5 - pig
ORGANISM #formal name Sus scrofa domestica #common name domestic pig
DATE 10-Apr-1996 #sequence_revision 24-May-1996 #text_change 03-Jun-1996
ACCESSION JC4584
REFERENCE White, M.E.; Diao, R.; Hathaway, M.R.; Mickelson, J.; Dayton, W.R.
#authors Biochem. Biophys. Res. Commun. (1996) 218:248-253
#journal Molecular cloning and sequence analysis of the porcine
#title insulin-like growth factor binding protein-5 complementary
#cross-references GB:U41340
#accession JC4584
#molecule_type mRNA
#residues 1-271 ##label WHI
#cross-references GB:U41340
#experimental source skeletal muscle
COMMENT This protein has essential roles in the regulation and coordination
of insulin-like growth factors action. This protein enhances the
in vitro activity of the insulin-like growth factors, plays a
role during myoblast proliferation and differentiation, and is
important in the growth and development of muscle tissue.
differentiation; growth factor; skeletal muscle
KEYWORDS #domain signal sequence #status predicted #label SIG\
1-19 #product insulin-like growth factor binding protein-5
FEATURE #status predicted #label MAT
20-271
SUMMARY #length 271 #molecular-weight 30323 #checksum 809
Query Match 4.7%; Score 128; DB 13; Length 271;
Best Local Similarity 37.5%; Pred. No. 8.02e-06;
Matches 30; Conservative 11; Mismatches 31; Indels 8; Gaps 7;
Db 2 vitavllllaacagpqlgsvhpcdekalmcpplgcclvkdpgccmcala 61
Qy 9 VVAVFVWLLALCSRPVAVGQNSGFC-RCPEPAPRC-PA-GVSLVLD-GCGCCRCVCAKQ 63

```

```

Db 62 eggscgvtyterc-aq-glrc 79
Qy 64 LGELCTE-RDPCDFPHKGLFC 82

RESULT 11
ENTRY JC1463 #type complete
TITLE insulin-like growth factor-binding protein 5 precursor - rat
ORGANISM #formal name Rattus norvegicus #common name Norway rat
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-1995
ACCESSION JC1463; A40403; F40403
REFERENCE JC1463
#authors Zhu, X.; Ling, N.; Shimaaki, S.
#journal Biochem. Biophys. Res. Commun. (1993) 190:1045-1052
#title Cloning of the rat insulin-like growth factor binding
protein-5 gene and DNA sequence analysis of its promoter
region.
#accession JC1463
#molecule_type DNA
#residues 1-271 ##label ZHU
#cross-references GB:L08275
REFERENCE A40403
#authors Shimaaki, S.; Shimonaka, M.; Zhang, H.P.; Ling, N.
#journal J. Biol. Chem. (1991) 266:10646-10653
#title Identification of five different insulin-like growth factor
binding proteins (IGFBPs) from adult rat serum and
molecular cloning of a novel IGFBP-5 in rat and human.
#cross-references MIM:9124847
#accession A40403
#molecule_type mRNA
#residues 1-271 ##label SHI
#cross-references GB:M62781
#accession F40403
#status preliminary
#molecule_type protein
#residues 20-253, 'X', 27-28, 'X', 30-36, 'X', 38-43, 'X', 45-51, 'X', 53,
'XX' ##label SH2

GENETICS
#introns 112/3; 188/3; 228/3
CLASSIFICATION #superfamily thyroglobulin type I repeat homology
FEATURE
1-19 #domain signal sequence #status predicted #label SIG\
20-271 #product insulin-like growth factor binding protein 5
#status predicted #label MAT\
191-262 #domain thyroglobulin type I repeat homology #label THY1
SUMMARY #length 271 #molecular-weight 30298 #checksum 1075
Query Match 4.6%; Score 125; DB 14; Length 271;
Best Local Similarity 36.0%; Pred. No. 2.33e-05;
Matches 27; Conservative 12; Mismatches 28; Indels 8; Gaps 7;
Db 7 llllaacavpqlgsvhpcdekalmcpplgcclvkepgccmcalaegqsc 66
Qy 14 VVLLALCSRPVAVGQNSGFC-RCPEPAPRC-PA-GVSLVLD-GCGCCRCVCAKQLGELC 68
Db 67 gvyterc-aq-glrc 79
Qy 69 TE-RDPCDFPHKGLFC 82

```

```
RESULT 12
ENTRY   JP0076      #type complete
TITLE   nel protein - chicken
ORGANISM  Gallus gallus #common name chicken
DATE     24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change
                                01-Dec-1995
ACCESSIONS A38963; JP0076
REFERENCE   A38963
#authors   Matsuhashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.;
            Taniguchi, S.; Hori, K.
#journal   Dev. Dyn. (1995) 203:212-222
#title     New gene, nel, encoding a Mr 93K protein with EGF-like
            repeats is strongly expressed in neural tissues of early
            stage chick embryos.
#accession A38963
#molecule_type mRNA
#residues 1-835 #label MAT
#cross-references DBJ:D45365
#experimental_source 9-day embryo
REFERENCE   JP0076
#authors   Matsuhashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.;
            Taniguchi, S.; Hori, K.
#submission submitted to JIPID, January 1995
#description A new gene, nel, encoding a Mr 93K protein with EGF-like
            repeats is strongly expressed in neural tissues of early
            stage chick embryos.
#accession JP0076
#molecule_type mRNA
#residues 1-835 #label MA2
#cross-references DBJ:D45365
#experimental_source 9-day embryo
CLASSIFICATION #superfamily von Willebrand factor type C repeat homology
FEATURE
273-333      #domain von Willebrand factor type C repeat homology
              #label VMCV
395-592      #region EGF-like repeats
SUMMARY      #length 835 #molecular-weight 93411 #checksum 7565
              4.6%; Score 126; DB 13; Length 835;
              Beat Local Similarity 40.3%; Pred. No. 1.63e-05;
              Matches 25; Conservative 10; Mismatches 20; Indels 7; Gaps 6;

Db 274 ctmkgmttyrefeswtgdgk-nctcmngtvqcealcalsdc-p-pnsalay-vd--gkcc 327
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 103 CIFGTVYRSGESFQSSCKYQCTCLDGAVGCMPL-CSMDVRLFSDFCPFRVRKLPKGC 161
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 328 ke 329
|
QY 162 EE 163

Query Match 4.6%; Score 126; DB 13; Length 835;
Beat Local Similarity 40.3%; Pred. No. 1.63e-05;
Matches 25; Conservative 10; Mismatches 20; Indels 7; Gaps 6;

Db 274 ctmkgmttyrefeswtgdgk-nctcmngtvqcealcalsdc-p-pnsalay-vd--gkcc 327
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 103 CIFGTVYRSGESFQSSCKYQCTCLDGAVGCMPL-CSMDVRLFSDFCPFRVRKLPKGC 161
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 328 ke 329
|
QY 162 EE 163

RESULT 13
ENTRY   B37252      #type complete
TITLE   insulin-like growth factor-binding protein 4 precursor -
            human
ALTERNATE_NAMES IGFBP-4; inhibitory insulin-like growth factor-binding
            protein; insulin-like growth factor-binding protein PC3
ORGANISM #formal name Homo sapiens #common name man
DATE     28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change
```

```
06-Sep-1996
ACCESSIONS B37252; B39842; A36549; A60712; A54650; A49801; A34419
REFERENCE   A37252
#authors   Shimasaki, S.; Uchiyama, F.; Shimonaka, M.; Ling, N.
#journal   Mol. Endocrinol. (1990) 4:1451-1458
#title     Molecular cloning of the cDNAs encoding a novel insulin-like
            growth factor-binding protein from rat and human.
#cross-references MUID:91133415
#accession B37252
#status     not compared with conceptual translation
#molecule_type mRNA
#residues 1-258 #label SHI
REFERENCE   A39842
#authors   Kiefer, M.C.; Masiarz, F.R.; Bauer, D.M.; Zapf, J.
#journal   J. Biol. Chem. (1991) 266:9043-9049
#title     Identification and molecular cloning of two new 30-kDa
            insulin-like growth factor binding proteins isolated from
            adult human serum.
#cross-references MUID:91225006
#accession B39842
#status     not compared with conceptual translation
#molecule_type mRNA
#residues 1-258 #label KIE
REFERENCE   A36549
#authors   Lafour, D.; Mohan, S.; Linkhart, T.A.; Baylink, D.J.; Strong,
            D.D.
#journal   Mol. Endocrinol. (1990) 4:1806-1814
#title     Inhibitory insulin-like growth factor-binding protein:
            cloning, complete sequence, and physiological regulation.
#cross-references MUID:91186988
#accession A36549
#status     not compared with conceptual translation
#molecule_type mRNA
#residues 1-50, 'A', 52-197, 'F', 199-258 #label LAT
REFERENCE   A60712
#authors   Perkel, V.S.; Mohan, S.; Baylink, D.J.; Linkhart, T.A.
#journal   J. Clin. Endocrinol. Metab. (1990) 71:533-535
#title     An inhibitory insulin-like growth factor binding protein
            (In-IGFBP) from human prostatic cell conditioned medium
            reveals N-terminal sequence identity with bone derived
            In-IGFBP.
#accession A60712
#status     not compared with conceptual translation
#molecule_type protein
#residues 22-26, 'X', 28-29, 'X', 31-35 #label PER
REFERENCE   A54650
#authors   Mohan, S.; Baylink, D.J.
#journal   Growth Regul. (1991) 1:110-118
#title     Evidence that the inhibition of TE85 human bone cell
            proliferation by agents which stimulate cAMP production may
            in part be mediated by changes in the IGF-II regulatory
            system.
#cross-references MUID:93091814
#accession A54650
#status     not compared with conceptual translation
#molecule_type protein
#residues 22-29, 'X', 31-42 #label MOH
#cross-references NCBI:P121076
#experimental_source TE85 osteosarcoma cells
#note       sequence extracted from NCBI backbone
REFERENCE   A49801
#authors   Culouscou, J.M.; Shoyab, M.
#journal   Cancer Res. (1991) 51:2813-2819
#title     Purification of a colon cancer cell growth inhibitor and its
```

identification as an insulin-like growth factor binding protein.

```
#accession A49801
#molecule_type protein
#residues_ 22-53 #label CUL
#experimental_source colon adenocarcinoma cells
REFERENCE A3419
#authors Mohan, S.; Bautista, C.M.; Wergedal, J.; Baylink, D.J.
#journal Proc. Natl. Acad. Sci. U.S.A. (1989) 86:8338-8342
#title Isolation of an inhibitory insulin-like growth factor (IGF)
binding protein from bone cell-conditioned medium: a
potential local regulator of IGF action.
#cross-references MUID:90046792
#accession A34419
#molecule_type protein
#residues_ 22-29, 'E', 31-32, 'A', 34-36 #label MO2
GENETICS GDB:IGFBP4
#gene GDB:IGFBP4
#cross-references GDB:126811
#map_position 17q12-17q21.1
CLASSIFICATION #superfamily thyroglobulin type I repeat homology
KEYWORDS glycoprotein
FEATURE
1-21 #domain signal sequence #status experimental #label SIG\
22-258 #product insulin-like growth factor-binding protein 4
#status experimental #label MAT\
174-249 #domain thyroglobulin type I repeat homology #label
THY1\
125 #binding site carbohydrate (Asn) (covalent) #status
experimental
SUMMARY #length 258 #molecular-weight 27934 #checksum 9064
Query Match 4.5%; Score 123; DB 13; Length 258;
Best Local Similarity 38.2%; Pred. No. 4.71e-05;
Matches 26; Conservative 14; Mismatches 22; Indels 6; Gaps 6;
Db 1 mlpclvaallaaagpgpslgde-aihcpcseeklarcrpvgceelvrepqgcgcate 59
| | : | | | : | | | : | | : | | | : | | : | | | : |
Qy 6 MGPVRVAFWLLALCSRAVGQSGPCR-CPDEPAPRC-PA-GVS-LVLD-GCGGCRVC 60
| | : | | | : | | | : | | : | | | : | | : | | | : |
Db 60 alglqmpc 67
| | | |
Qy 61 AKQLGELC 68
RESULT 14
ENTRY #type complete
TITLE TGF-beta-related protein 60A precursor - fruit fly
ALTERNATE_NAMES (Drosophila melanogaster)
ORGANISM bone morphogenetic protein homolog precursor
#formal_name Drosophila melanogaster
DATE 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change
10-Nov-1995
ACCESSIONS A43918; A41233
REFERENCE A43918
#authors Doctor, J.S.; Jackson, P.D.; Rashka, K.E.; Visalli, M.;
Hoffmann, F.M.
#journal Dev. Biol. (1992) 151:491-505
#title Sequence, biochemical characterization, and developmental
expression of a new member of the TGF-beta superfamily in
```

Drosophila melanogaster.

```
#cross-references MUID:92290120
#accession A43918
#status preliminary
#molecule_type mRNA
#residues_ 1-455 #label DOC
#cross-references NCBI:106399; NCBI:106400
#note sequence extracted from NCBI backbone
REFERENCE A41233
#authors Wharton, K.A.; Thomsen, G.H.; Gelbart, W.M.
#journal Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9214-9218
#title Drosophila 60A gene, another transforming growth factor beta
family member, is closely related to human bone
morphogenetic proteins.
#cross-references MUID:92021021
#accession A41233
#molecule_type mRNA
#residues_ 1-455 #label WHA
#cross-references GB:M77012
COMMENT This protein is a member of the transforming growth factor beta
family.
GENETICS 60A
#gene 60A
CLASSIFICATION #superfamily inhibin
KEYWORDS homodimer
SUMMARY #length 455 #molecular-weight 51687 #checksum 2124
Query Match 4.5%; Score 121; DB 6; Length 455;
Best Local Similarity 43.2%; Pred. No. 9.49e-05;
Matches 16; Conservative 9; Mismatches 10; Indels 2; Gaps 2;
Db 419 ccaptrlgalpvlyhndenvnlkkyrmimvkschg 455
| | : | | | : | | | : | | : | | | : | | : | | | : |
Qy 292 CCTPHRTTTLPVFKCPDGEV-MKKNM-MFIKTCACH 326
| | : | | | : | | | : | | : | | | : | | : | | | : |
RESULT 15
ENTRY #type complete
TITLE insulin-like growth factor binding protein-4 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
02-Jul-1996
ACCESSIONS I48603
REFERENCE I48600
#authors Schuller, A.G.; Groffen, C.; van Neck, J.W.; Zwartthoff, F.
Drop, S.L.
#journal Mol. Cell. Endocrinol. (1994) 104:57-66
#title cDNA cloning and mRNA expression of the six mouse
insulin-like growth factor binding proteins.
#cross-references MUID:95121750
#accession I48603
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-254 #label RES
#cross-references EMBL:X81582; NID:g550382; CDS_PID:g550383
GENETICS gene name IGFBP-4
SUMMARY #length 254 #molecular-weight 27760 #checksum 4740
Query Match 4.3%; Score 116; DB 14; Length 254;
```

Search completed: Wed Sep 17 09:34:50 1997
Job time : 82 secs.

maryh@stic

4381-1

NeWSprinter20

Wed Sep 17 11:17:56 1997

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5

MPARCH_PP

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MParch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 17 09:32:15 1997; MaePar time 9.17 Seconds
Tabular output not generated. 807.003 Million cell updates/sec

Title: >NS-08-167-628-2
Description: (1-349) from 3408040.ppe
Perfect Score: 2713
Sequence: 1 MTAASMGPRVAVFVLLALC.....PGNDIFESLYRMYGDMA 349

Scoring table: PAM 150
Gap 11

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot34
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 44.133; Variance 66.349; scale 0.665

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2713	100.0	349	CTGF_HUMAN	CONNECTIVE TISSUE GRO	0.00e+00
2	2514	92.7	348	CTGF_MOUSE	CONNECTIVE TISSUE GRO	0.00e+00
3	1478	54.5	343	NOV_XENLA	NOV PROTEIN HOMOLOG P	0.00e+00
4	1460	53.8	351	NOV_CHICK	NOV PROTEIN PRECURSOR	0.00e+00
5	1450	53.4	353	NOV_COTJA	NOV PROTEIN PRECURSOR	0.00e+00
6	1415	52.2	357	NOV_HUMAN	NOV PROTEIN HOMOLOG P	0.00e+00
7	1152	42.5	375	CELO_CHICK	CEF-10 PROTEIN PRECUR	6.10e-277
8	1113	41.0	379	CYR6_MOUSE	CYR61 PROTEIN PRECURS	4.24e-266
9	133	4.9	254	IBP4_RAT	INSULIN-LIKE GROWTH F	5.18e-09

10	128	4.7	867	9	SSPO_BOVIN	SCO-SPONDIN (FRAGMENT	4.95e-08
11	125	4.6	271	5	IBP5_RAT	INSULIN-LIKE GROWTH F	1.88e-07
12	123	4.5	258	5	IBP4_HUMAN	INSULIN-LIKE GROWTH F	4.54e-07
13	121	4.5	455	1	GOA_DROME	GOA PROTEIN PRECURSOR	1.09e-06
14	116	4.3	254	5	IBP4_MOUSE	INSULIN-LIKE GROWTH F	9.35e-06
15	113	4.2	272	5	IBP1_RAT	INSULIN-LIKE GROWTH F	3.32e-05
16	113	4.2	2813	10	VWF_HUMAN	VON WILLEBRAND FACTOR	3.32e-05
17	110	4.1	263	5	IBP1_BOVIN	INSULIN-LIKE GROWTH F	1.16e-04
18	111	4.1	271	5	IBP5_MOUSE	INSULIN-LIKE GROWTH F	7.65e-05
19	111	4.1	304	5	IBP2_RAT	INSULIN-LIKE GROWTH F	7.65e-05
20	111	4.1	305	5	IBP2_MOUSE	INSULIN-LIKE GROWTH F	7.65e-05
21	110	4.1	2274	6	MUC2_HUMAN	MUCIN 2 (INTESTINAL M	1.16e-04
22	109	4.0	272	5	IBP1_MOUSE	INSULIN-LIKE GROWTH F	1.75e-04
23	106	3.9	258	5	IBP4_BOVIN	INSULIN-LIKE GROWTH F	5.92e-04
24	105	3.9	272	5	IBP5_HUMAN	INSULIN-LIKE GROWTH F	8.86e-04
25	106	3.9	807	4	FSPO_RAT	F-SPONDIN PRECURSOR	5.92e-04
26	102	3.8	259	5	IBP1_HUMAN	INSULIN-LIKE GROWTH F	2.92e-03
27	102	3.8	291	5	IBP3_BOVIN	INSULIN-LIKE GROWTH F	2.92e-03
28	104	3.8	311	5	IBP2_CHICK	INSULIN-LIKE GROWTH F	1.32e-03
29	103	3.8	328	5	IBP2_HUMAN	INSULIN-LIKE GROWTH F	1.97e-03
30	103	3.8	500	10	VLI_HPV6B	MAJOR CAPSID PROTEIN	1.97e-03
31	102	3.8	798	5	ITB7_HUMAN	INTEGRIN BETA-7 SUBUN	2.92e-03
32	104	3.8	3133	5	HMCT_BOMMO	HEMOCYTIN PRECURSOR	1.32e-03
33	101	3.7	388	3	CSP_FLARE	CIRCUMSPOROZOITE PROT	4.32e-03
34	101	3.7	399	1	BM6A_MOUSE	BONE MORPHOGENETIC PR	4.32e-03
35	101	3.7	402	1	BMP6_HUMAN	BONE MORPHOGENETIC PR	4.32e-03
36	101	3.7	405	2	CSP_PLAFO	CIRCUMSPOROZOITE PROT	4.32e-03
37	101	3.7	412	2	CSP_PLAFA	CIRCUMSPOROZOITE PROT	4.32e-03
38	101	3.7	424	3	CSP_PLAFT	CIRCUMSPOROZOITE PROT	4.32e-03
39	100	3.7	442	3	CSP_PLAFW	CIRCUMSPOROZOITE PROT	6.38e-03
40	101	3.7	551	11	YVD3_CAEEL	HYPOTHETICAL ZINC MET	4.32e-03
41	99	3.6	111	5	IBP5_BOVIN	INSULIN-LIKE GROWTH F	9.40e-03
42	99	3.6	291	5	IBP3_MOUSE	INSULIN-LIKE GROWTH F	9.40e-03
43	99	3.6	292	5	IBP3_RAT	INSULIN-LIKE GROWTH F	9.40e-03
44	98	3.6	806	5	ITB7_MOUSE	INTEGRIN BETA-7 SUBUN	1.38e-02
45	96	3.5	803	4	FSPO_XENLA	F-SPONDIN PRECURSOR	2.95e-02

ALIGNMENTS

RESULT	1	CTGF_HUMAN	STANDARD;	PRT;	349 AA.
ID	CTGF_HUMAN				
AC	P29279;				
DT	01-DEC-1992 (REL. 24, CREATED)				
DT	01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)				
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)				
DE	CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.				
GN	CTGF.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;				
RX	MEDLINE; 91373462.				
RA	BRADHAM D.M., IGARASHI A., POTTER R.L., GROTEENDORST G.R.;				
RL	J. CELL BIOL. 114:1285-1294 (1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;				

RX	MEDLINE; 93187114.
RA	IGARASHI A., BRADHAM D.M., OKOCHI H., GROTDENDORST G.R.;
RJ	J. DERMATOL. 19:642-643(1992).
[3]	
RP	SEQUENCE FROM N.A.
RR	OEMAR B.S., WERNER A., YANG Z., GARNIER J.M., GENTZ R., LUESCHER T.F.;
RA	SUBMITTED (APR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC	-!- FUNCTION: MAJOR CONNECTIVE TISSUE MITOTACTICANT SECRETED BY
CC	HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN
CC	MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.
CC	-!- SUBUNIT: MONOMER.
CC	-!- ALTERNATIVE PRODUCTS: A SHORTER FORM MAY BE PRODUCED BY
CC	ALTERNATIVE SPLICING OF THE SAME GENE.
CC	-!- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
CC	TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
CC	-!- SIMILARITY: CONTAINS A VMFC DOMAIN.
CC	-!- SIMILARITY: CONTAINS A C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).
DR	EMBL; M92934; G180924; -;
DR	EMBL; S56201; G268235; -;
DR	EMBL; X78947; G474934; -;
DR	PIR; A40551; A40551.
DR	PIR; S44205; S44205.
DR	MIM; 121009; -;
DR	PROSITE; PS00222; IGF BINDING.
KW	GROWTH FACTOR BINDING; SIGNAL; ALTERNATIVE SPLICING.
FT	SIGNAL 1 21
FT	CHAIN 22 349
FT	DOMAIN 101 167
FT	DOMAIN 256 330
FT	D1SULFID 256 293 BY SIMILARITY.
FT	D1SULFID 273 307 BY SIMILARITY.
FT	D1SULFID 284 323 BY SIMILARITY.
FT	D1SULFID 287 325 BY SIMILARITY.
FT	D1SULFID 292 329 BY SIMILARITY.
FT	CARBOHYD 28 28 POTENTIAL.
FT	CARBOHYD 225 225 POTENTIAL.
FT	VARSPLIC 172 198 MISSING (LN SHORT FORM).
SEQ	SEQUENCE 349 AA; 38069 MW; C21E9662 CRC32;
Query Match 100.0%; Score 2713; DB 3; Length 349;	
Best Local Similarity 100.0%; Pred. No. 0.00e+00;	
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps	
Db	1 mtaasgprvrvafvllalcrspavqncgpcrcdpapropagvalvdggccrvc 60
Qy	1 MTAASMGPRVRVAVFWLLALCSRPAVGQCSCGRCPCDEPAPRCAGVLVDGGCCRCVC 60
Db	61 akqlgelcterdpcdphkglfcdfgsparnkigvctaakdgapciffgtvyrasefgesc 120
Qy	61 AKQGELECTERDCDPHKGLFCDFGSANRKLGICTAKOGAPCIFGCTVYRSGESFQSSC 120
Db	121 kyqtclidgavgcmplcsmdvrlpsdcpffrrvk lpgkceewwcdpkdqtvvgpala 180
Qy	121 KYQCTCLDGAVGCMPLCSMDVRLPSDCPPFRVRKLPKCCEEWCDPKDQTVVGPALA 180
Db	181 ayrledtfgdpmritranclvdtwsacsktcgmgi strvtndnasrlekqarlcnvr 240
Qy	181 AYRLEDTFGDDPTMIRANCLVDTWSACSKTCGMGI STRVTNDNASRLEKQSRLCNVR 240
Db	241 pceadleenikkakkcilrtbpkiekpbikfelqsctemktvtrakfcgvctdgrcctphrttt 300

[illegible]

Qy	1	MTAASGQPVRAVFVLLALCSRAVAGQNSGCEPCDEPAPCPAGVSLVLDGCGCQRCV	60
Db	60	akqlgelcterdpdphkgilfcagfepanrk.igvtakdgapcvfgsgvyrsgesfqsc	119
Qy	61	AKQLGELCTERDPDOPKHGILCFDGFSPANRKIGVCTAKDGAFCIFGFTVYRSGESFQSC	120
Db	120	kyqctcldgavgcwplcemdvrlpsdpcefprvrlpqkckewcdcepdkdrtaavgpala	179
Qy	121	KYQCTCLDGAVGCMPLCSMDVRLPSDCEFPFRVLKPKCCCEEWVCDPEKDKQIVVGPALA	180
Db	180	ayrledtfgpdpmtmrancivqtwtwsacskctgmgiatrvntndntflekgsrlcmvr	239
Qy	181	AYRLEDTFGPDPMTIRANCIQVQTEWSACSKCTGCMGISTRTVNDNASCRLKQSRLCMVR	240
Db	240	pceadleenikkgkclrtbpkiakpvkf.legctavkytrakfgcvctdgrctphrttt	299
Qy	241	PCEAADLEENIKKGKCLRTPKI SKP IKFELSGCTSKWITYRAKFGCVCTDGRCTPHRTTT	300
Db	300	lpvefkcpdgeimkmmnfiktcachyncpgdndifaelyyrkygmga	348
v	301	LPVEFKCPDGEVMMKNMNFIKTCACHYNCPGDNDFESLIVRYKMGDMA	349

[illegible]

Query Match 54.5%; Score 1478; DB 6; Length 343;
Best Local Similarity 53.7%; Pred. No. 0.00e+00;
Matches 180; Conservative 69; Mismatches 78; Indels 8; Gaps 8;

Db	1	mtp-hlalcfil-l-igqvaskcpseqdcqpeep-pecapavllildgcgcpvcarg	56
Qy	6	MGFVRVFWLLALCSRPVAGVCGSGFC-RCPEPAPRPAGVSLVLDGCGCCRCVAKQL	64
Db	57	geschlnpcedekglycefnadprmetgtcmalegnacvfdgvvyrnresfpackyh	116
Qy	65	GELCTERPDCPHKGLFCDFGSPANRKIGVCAKAGAPCIFGGTVYRSGSFOSCKYQC	124
Db	117	tclnghicvprcnldlllpqpcprrrvkvpgecekwwcd-akeemalggfamaavr	175
Qy	125	TCLDGAVGMPCLCSMDVRLPSPDCPPFRVRKPLGKCEEWDEPKQTVVGP-ALAAVR	183
Db	176	peatigidasstsfaciagqtewsacsctcmgvasrvtnarctcmgkqirlcmvrace	235
Qy	184	LEDTGPDPTIRANCLVQITWSACSKTCGMGISTRVNDNASCLREKQSRUCMRPCE	247
Db	236	eepgwvhvekkqkvrvrktkplhfyknctsvqpkpkfcgcsdgdroctphstktmh	293
Qy	244	ADLEENI-KGKKKCI RTPKISKPIKFELSGCTSMKTYRAKFCGCTDGRCTPRTTLP	302
Db	296	vefvcpqkriwkpvmvistcvchynpcdsallq	330
Qy	303	VEFKFGDGEVMKKGNMFIRKTCACHYCPGDNDFIE	337

[illegible]

DR PIR; S20078; S20078.
 DR PROSITE; PS00222; IGF BINDING.
 KW PROTO-ONCOGENE; GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 351 NOV PROTEIN.
 FT DOMAIN 104 170 VWFC.
 FT DOMAIN 258 332 CTCK.
 FT DISULFID 258 295 BY SIMILARITY.
 FT DISULFID 275 309 BY SIMILARITY.
 FT DISULFID 286 325 BY SIMILARITY.
 FT DISULFID 289 327 BY SIMILARITY.
 FT DISULFID 294 331 BY SIMILARITY.
 FT CARBOHYD 274 274 POTENTIAL.
 SQ SEQUENCE 351 AA; 38268 MW; C7044065 CRC32;
 Query Match 53.8%; Score 1460; DB 6; Length 351;
 Best Local Similarity 54.2%; Pred. No. 0.00e+00;
 Matches 186; Conservative 69; Mismatches 80; Indels 8; Gaps 7;
 Db 3 tggggglpvlvllllllrpcevsgraeaprcpgrcgaep-prcapgvavldgocgel 61
 Qy 2 TAASNG-PVRVAFVLLALCSRAVQNGSGPC--RCPDEPAPRCAGVSLVLDGCGCCR 58
 Db 62 vcargrseapllpcdesaglycdrgedgggagicmvlcdncvfdgmiyrngctfcp 121
 Qy 59 VCAKQLGELCTERDPCDPHGLFCDFGSPANRKGVCCTAKDGAPCFGGTVYRSGESFOS 118
 Db 122 eckvctcdqgiclpncnlglllpgdpdpfrkievpgcecekwcd-prdevillgcf 180
 Qy 119 SKQYCTCLDGAVGCMPLCSMDVRLSPDCFPFRVKLPKCCEEWCDPEKQTVVGP- 177
 Db 181 amayrqaetlgidvdsasancieqttesasckscgmgtfstrvnrngqcmvktqrlc 240
 Qy 178 ALAAVRLDFTGPDFTMRANCLVQTTEWSACSKTCGMGISTRTVNDNASCRLEKQSLC 237
 Db 241 mmrpeene-epsdkkgkckitckkkmkavrfeykntcsavtqkprycglndgrecctphn 299
 Qy 238 WVRPEADLEENIKKGGKICIRTPKISKPILFSGCTSMKTYRAKFGVCTDGRCTPHR 297
 Db 300 ttiqvefrcpgkflkkpmlintevchncpqsnnaffcp 342
 Qy 298 TTTLPVEKPDGEVKKNNMFIKTACHYNCP-GNDIFESL 339
 RESULT 5
 ID NOV COTJA STANDARD; PRT; 353 AA.
 AC P42642;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE NOV PROTEIN PRECURSOR.
 GN NOV.
 OS COTURNIX COTURNIX JAPONICA (JAPANESE QUAIL).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
 OC GALLIFORMES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA WEISKIRCHEN R., BISTER K.;
 RL SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL

CC GROWTH REGULATION (BY SIMILARITY).
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FLSP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A VWFC DOMAIN.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; U13063; G532697; -.
 DR PROSITE; PS00222; IGF BINDING.
 KW PROTO-ONCOGENE; GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 353 NOV PROTEIN.
 FT DOMAIN 106 172 VWFC.
 FT DOMAIN 260 334 CTCK.
 FT DISULFID 260 297 BY SIMILARITY.
 FT DISULFID 277 311 BY SIMILARITY.
 FT DISULFID 288 327 BY SIMILARITY.
 FT DISULFID 291 329 BY SIMILARITY.
 FT DISULFID 296 333 BY SIMILARITY.
 FT CARBOHYD 276 276 POTENTIAL.
 SQ SEQUENCE 353 AA; 38667 MW; C4F5928D CRC32;
 Query Match 53.4%; Score 1450; DB 6; Length 353;
 Best Local Similarity 54.7%; Pred. No. 0.00e+00;
 Matches 185; Conservative 66; Mismatches 78; Indels 9; Gaps 8;
 Db 10 pvlvllllllllrpcevsgraeaprcpgrcgaep-prcapgvavldgocgelvcarq 68
 Qy 8 PVRVAFVLLALCSRAVQNGSGPC--RCPDEPAPRCAGVSLVLDGCGCCRCAQ 63
 Db 69 rgescapllpcdesaglycdrgedgggagicmvlcdncvfdgmiyrngctfcpckcyq 128
 Qy 64 LSELCTERDPCDPHGLFCDFGSPANRKGVCCTAKDGAPCFGGTVYRSGESFOSCKYQ 123
 Db 129 ctcdqgiclpncnlglllpgdpdpfrkievpgcecekwcd-prdevillgcfamaay 187
 Qy 124 CTCLDGAVGCMPLCSMDVRLSPDCFPFRVKLPKCCEEWCDPEKQTVVGP-ALAA 182
 Db 188 rqaetlgidvdsasancieqttesasckscgmgtfstrvnrngqcmvktqrlcmmrpe 247
 Qy 183 RLEDFTGPDFTMRANCLVQTTEWSACSKTCGMGISTRTVNDNASCRLEKQSLCVRPC 242
 Db 248 ene-epsdkkgkckitckkkmkavrfeykntcsavtqkprycglndgrecctphntkiq 306
 Qy 243 EADLEENIKKGGKICIRTPKISKPILFSGCTSMKTYRAKFGVCTDGRCTPHRTTLP 302
 Db 307 vefrcpgkflkkpmlintevchncpqsnnaffcp 344
 Qy 303 VEFKPDGEVKKNNMFIKTACHYNCP-GNDIFESL 339
 RESULT 6
 ID NOV HUMAN STANDARD; PRT; 357 AA.
 AC P48745;
 DT 01-FEB-1996 (REL. 33, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE NOV PROTEIN HOMOLOG PRECURSOR (NOVH).
 GN NOV.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.

Listing for Mary Hale Wed Sep 17 17:06:07 1997

Db 245 rvtndpckliketricevtrpgpsvaalkgkktktpspvrfyagssvsky 304
QY 220 RVTNDNASCRLEKQSRCLMWPCPEADLEENIKKGGKCIKIRTPKISKPIKFLSGCTSMKY 279
Db 305 rpkyscscvdcgctcpqtrvkirfcdgctfukvmmiqscrcnynpcphaneay-p- 362
QY 280 RAKFGVCTDGRCCPHRTTILPVEFKCPDGEVKKMMFMFKTACACHNCPGNDIFESL 339
Db 363 fvr 365
QY 340 YVR 342

RESULT 8
ID CYR6 MOUSE STANDARD; PRT; 379 AA.
AC P18406;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CYR61 PROTEIN PRECURSOR (3CH61).
GN CYR61.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C / 3T3;
RX MEDLINE: 90287146.
RA O'BRIEN T.P., YANG G.P., SANDERS L., LAU L.F.;
RL MOL. CELL. BIOL. 10:3569-3577(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AJ; TISSUE=EMBRYONIC FIBROBLAST;
RX MEDLINE: 91288203.
RA LATINKIC B.V., O'BRIEN T.P., LAU L.F.;
RL NUCLEIC ACIDS' RES. 19:3261-3267(1991).
CC -!- FUNCTION: CYR61 MIGHT ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
CC PROTEINS.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN
CC NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
CC -!- INDUCTION: BY GROWTH FACTORS.
CC -!- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
CC AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST
CC IN LUNG.
CC -!- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
CC -!- SIMILARITY: CONTAINS A VWFC DOMAIN.
CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
DR EMBL; M32490; G309206; -.
DR EMBL; X56790; G50633; -.
DR PIR; A35669; A35669.
DR PROSITE; PS00222; IGF BINDING.
KW GROWTH FACTOR BINDING; SIGNAL.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 379 CYR61 PROTEIN.
FT DOMAIN 98 164 VWFC.
FT DOMAIN 284 358 CTCK.
FT DISULFID 284 321 BY SIMILARITY.
FT DISULFID 301 335 BY SIMILARITY.
FT DISULFID 312 351 BY SIMILARITY.

Listing for Mary Hale Wed Sep 17 17:16:07 1997

FT DISULFID 315 353 BY SIMILARITY.
FT DISULFID 320 357 BY SIMILARITY.
SQ SEQUENCE 379 AA; 41709 MW; 116880C7 CRC32;
Query Match 41.0%; Score 1113; DB 3; Length 379;
Best Local Similarity 48.3%; Pred. No. 4,24e-266;
Matches 180; Conservative 79; Mismatches 76; Indels 38; Gaps 20;
Db 1 masstfrtlava-vtllhl-triala-tcpaachople-apkcapgvglvrdgcckvc 56
QY 1 MTAASMGVRAVAFVLLALCSRPAPVQNGSGPCRCPCDEPAPRCAPAGVSLVLDGCGCCRV 60
Db 57 akqlnedcaktqcdhtkglecnfgasatalkgicraqegprcpcnarynryngesfcpn 116
QY 61 AKQLGELCTERDPCPHKGLFCDFGSPANRKGVCCTAK-DGAPCIFGGTVYRSGSFQSS 119
Db 117 ckhqctcidgavcplcpqelalpnlgcpnprlvkvsqccceewcdedsikdldq 176
QY 120 CKYQCTCLDAGVGMPLCSMDVRLPSPDCPFPRRVKLPKCCCEEWVCE-P-KD---QT 173
Db 177 dlilgdaseveltrnnelliaigkgsalkrlpvfgteprvifnplhahgqciavttsaq 236
QY 174 -WVG--PA---LA-----AY-R---LE--DTFGDP--TM---IRAN---CIVQTTE 208
Db 237 cskactgtgistrvtnndnpeclrvketricvrcpgpvyasalkgkksktskspvrf 296
QY 209 CSKTCGMGISTRTVNDNASCRLEKQSRCLMWPCPEADLEENIKKGGKCIKIRTPKISKPIK 268
Db 297 tyagcavkkyrpkyscscvdcgctcpqtrvkirfcdgctfukvmmiqscrcnyn 356
QY 269 ELSGTSMKYTRAKFGVCTDGRCCPHRTTILPVEFKCPDGEVKKMMFMFKTACACHN 328
Db 357 cphneasf-rly 368
QY 329 CPDNDI-FESLY 340

RESULT 9
ID IBP4 RAT STANDARD; PRT; 254 AA.
AC P21744;
DT 01-MAY-1991 (REL. 18, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 PRECURSOR (IGFBP-4)
DE (IBP-4) (IGF-BINDING PROTEIN 4).
GN IGFBP4 OR IGFBP-4.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91133415.
RA SHIMASAKI S., UCHIYAMA F., SHIMONAKA M., LING N.;
RL MOL. ENDOCRINOL. 4:1451-1458(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY;
RC MEDLINE: 93176147.
RA GAO L., LING N., SHIMASAKI S.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 190:1053-1059(1993).

RN SEQUENCE OF 22-61.
 RC TISSUE=SERUM;
 RX MEDLINE; 90073708.
 RA SHIMONAKA M., SCHROEDER R., SHIMASAKI S., LING N.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 165:189-195(1989).
 CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC INTERACTION EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
 CC PROMOTING EFFECTS OF THE IGFs WITH THEIR CELL SURFACE RECEPTORS.
 CC -!- BINDS IGF-II MORE THAN IGF-I.
 CC -!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
 DR EMBL; L08276; E73337; -.
 DR PIR; B33570; B33570.
 DR PIR; A37252; A37252.
 DR PIR; J01464; J01464.
 DR PIR; E40403; E40403.
 DR PROSITE; PS00222; IGF BINDING.
 DR PROSITE; PS00484; THYROGLOBULIN 1.
 KW GROWTH FACTOR BINDING; SIGNAL; GLYCOPROTEIN.
 FT SIGNAL 1 21
 FT CHAIN 22 254 INSULIN-LIKE GROWTH FACTOR BINDING
 FT CARBOHYD 125 125 PROTEIN 4.
 FT DOMAIN 196 245 POTENTIAL.
 FT SEQUENCE 254 AA; 27745 MW; 355316DC CRC32;
 Query Match 4.9%; Score 133; DB 5; Length 254;
 Best Local Similarity 39.7%; Pred. No. 5.18e-09;
 Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6;
 Db 1 mlpfglaallaaagrpqlde-aihcpcseeklarcpvpgvgeelvrpgcgcatc 59
 QY 6 MGPVRVAVVLLALCSRAVQNGSGPCR-CPDEPAPRC-PA-GVS-LVLD-GCGGCRVC 60
 Db 60 alglgmpe 67
 QY 61 AKQLGELC 68
 RESULT 10
 ID SSP BOVIN STANDARD; PRT; 867 AA.
 AC P98167;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DE SCO-SPONDIN (FRAGMENT).
 OS BOS TAURUS (BOVINE).
 OC EUKARYOTA; ANIMALIA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA;
 OC THERIA; EUTHERIA; ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EPENDYMOCYTE;
 RX MEDLINE; 96338614.
 RA GORON S., MONNERIE H., MEINIEL R., CREVEAUX I., LEHMANN W.,
 RA LAMALLE D., DASTUGUE B., MEINIEL A.;
 RL J. CELL SCI. 109:1053-1061(1996).
 CC -!- FUNCTION: INVOLVED IN THE MODULATION OF NEURONAL AGGREGATION.
 CC -!- TISSUE SPECIFICITY: SUBCOMMISSURAL ORGAN.
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -!- DEVELOPMENTAL STAGE: EMBRYO.
 CC -!- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.
 CC -!- SIMILARITY: CONTAINS AT LEAST 2 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS AT LEAST 3 LDL-RECEPTOR CLASS A DOMAINS.
 DR EMBL; X93922; E213357; -.
 KW GLYCOPROTEIN; CELL ADHESION; CALCIUM-BINDING; REPEAT; EGF-LIKE DOMAIN.
 FT NON TER 1
 FT DOMAIN 103 142 EGF-LIKE.
 FT DOMAIN 143 180 EGF-LIKE.
 FT DOMAIN 506 544 LDL-RECEPTOR CLASS A.
 FT DOMAIN 663 701 LDL-RECEPTOR CLASS A.
 FT DOMAIN 723 761 LDL-RECEPTOR CLASS A.
 FT CARBOHYD 88 88 POTENTIAL.
 FT CARBOHYD 309 309 POTENTIAL.
 FT CARBOHYD 409 409 POTENTIAL.
 FT NON TER 867 867
 SQ SEQUENCE 867 AA; 91817 MW; B1224081 CRC32;
 Query Match 4.7%; Score 128; DB 9; Length 867;
 Best Local Similarity 28.3%; Pred. No. 4.95e-08;
 Matches 32; Conservative 21; Mismatches 48; Indels 12; Gaps 12;
 Db 81 cpxdgcgvntcagelyfhacvpcplctddisqgatepdr-pcggp-gcwcpgagqvlgaa 138
 QY 35 CPDEPAPRC-PAGVSLVLDGCGGCRV-CAKQLGEL-CT-ERDFCDPHKGLFCDFGSPANRK 91
 Db 139 -gcvvprg-cpcolvdgrywpqrvktcdql-cvcqdgprprcpsidcavn 188
 QY 92 IGVCT-AKDGAPCIGFTVYRSGESFOSSCKYQCTCLDG-AVGCMF-L-CSMD 140
 RESULT 11
 ID IBP5 RAT STANDARD; PRT; 271 AA.
 AC P24594;
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 5 PRECURSOR (IGFBP-5)
 DE (IBP-5) (IGF-BINDING PROTEIN 5).
 GN IGFBP5 OR IGFBP-5.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-53.
 RC TISSUE=OVARY;
 RX MEDLINE; 91244847.
 RA SHIMASAKI S., SHIMONAKA M., ZHANG H.-P., LING N.;
 RL J. BIOL. CHEM. 266:10646-10653(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE; 93176146.
 RA ZHU X., LING N., SHIMASAKI S.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 190:1045-1052(1993).
 CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
 CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
 CC -!- TISSUE SPECIFICITY: MOSTLY IN KIDNEY.

RA WHARTON K.A., THOMSEN G.H., GELBART W.M.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:9214-9218(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92290120.
 RA DOCTOR J.S., JACKSON P.D., RASHKA K.E., VISALLI M., HOFFMANN F.M.;
 RL DEV. BIOL. 151:491-505(1992)
 CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC OF TRANSCRIPTION DURING EARLY EMBRYOGENESIS, IN PUPAE, AND IN
 CC ADULT MALES.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.
 DR EMBL; M7012; G156728; -.
 DR EMBL; M84795; G156730; -.
 DR PIR; A41233; A41233.
 DR PIR; A43918; A43918.
 DR HSP; P08112; 1TFG.
 DR FLYBASE; FBGN004788; TGF-BETA-60A.
 DR PROSITE; PS00250; TGF BETA.
 KW GROWTH FACTOR; CYTOKINE; GLYCOPROTEIN; SIGNAL.
 FT SIGNAL 1 ?
 FT PROPEP ? 335
 FT CHAIN 336 455
 FT DISULFID 354 420
 FT DISULFID 383 452
 FT DISULFID 387 454
 FT DISULFID 419 419
 FT DISULFID 238 238
 FT CARBOHYD 250 250
 FT CARBOHYD 396 396
 FT CARBOHYD 435 AA; 51687 MW; 6FAAC7CD CRC32;
 SQ SEQUENCE 435 AA; 51687 MW; 6FAAC7CD CRC32;
 Query Match 4.5%; Score 121; DB 1; Length 455;
 Best Local Similarity 43.2%; Pred. No. 1.09e-06;
 Matches 16; Conservative 9; Mismatches 10; Indels 2; Gaps 2;
 Db 419 ccaptrigalpvlhndenvnlkkyrnmivkscgch 455
 ||:| | :|||: | :| :| :||:| :||
 Qy 292 CCTPRTTLEVEFKPDGEV-MKKNM-MFIKTCACH 326
 RESULT 14
 ID IBP4_MOUSE STANDARD; PRT; 254 AA.
 AC P47879;
 DT 01-FEB-1996 (REL. 33, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 PRECURSOR (IGFBP-4)
 DE (IBP-4) (IGF-BINDING PROTEIN 4).
 GN IGFBP4 OR IGFBP-4.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE; 95121750.
 RA SCHULLER A.G.P., GROFFEN C., VAN NECK J.W., ZWARTHOFF E.C.,
 RA DROP S.L.S.;
 RL MOL. CELL. ENDOCRINOL. 104:57-66(1994).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE; 95152444.
 RA BETHLE C.R., VITULLO J.C., MILLER R.E., ARON D.C.;
 RL BIOCHEM. MOL. BIOL. INT. 34:385-392(1994).
 CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC OF TRANSCRIPTION DURING EARLY EMBRYOGENESIS, IN PUPAE, AND IN
 CC ADULT MALES.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.
 DR EMBL; M7012; G156728; -.
 DR EMBL; M84795; G156730; -.
 DR PIR; A41233; A41233.
 DR PIR; A43918; A43918.
 DR HSP; P08112; 1TFG.
 DR FLYBASE; FBGN004788; TGF-BETA-60A.
 DR PROSITE; PS00250; TGF BETA.
 KW GROWTH FACTOR; CYTOKINE; GLYCOPROTEIN; SIGNAL.
 FT SIGNAL 1 ?
 FT PROPEP ? 335
 FT CHAIN 336 455
 FT DISULFID 354 420
 FT DISULFID 383 452
 FT DISULFID 387 454
 FT DISULFID 419 419
 FT DISULFID 238 238
 FT CARBOHYD 250 250
 FT CARBOHYD 396 396
 FT CARBOHYD 435 AA; 51687 MW; 6FAAC7CD CRC32;
 SQ SEQUENCE 435 AA; 51687 MW; 6FAAC7CD CRC32;
 Query Match 4.3%; Score 116; DB 5; Length 254;
 Best Local Similarity 36.9%; Pred. No. 9.35e-06;
 Matches 24; Conservative 14; Mismatches 21; Indels 6; Gaps 6;
 Db 1 mlpcslvaallltagrpalsgde-aihcpcseeklarcpvpgceelivrepqgcscatc 59
 ||:| | :|||: | :| :| :||:| :||
 Qy 6 MGEVRVAFVVVLLALCSRPVAVGQNGSGPCR-CPDEPAPRC-PA-GVS-IVLD-GCGCCRCVC 60
 ||:| | :|||: | :| :| :||:| :||
 Db 60 alqlg 64
 Qy 61 AKQLG 65
 RESULT 15
 ID IBP1_RAT STANDARD; PRT; 272 AA.
 AC P21743;
 DT 01-MAY-1991 (REL. 18, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (IGFBP-1)
 DE (IBP-1) (IGF-BINDING PROTEIN 1).
 GN IGFBP1 OR IGFBP-1.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=DECIDUA;
 RX MEDLINE; 90231347.
 RA MURPHY L.J., SENEVIRATNE C., BALLEJO G., CROZE F., KENNEDY T.G.;
 RA MOL. ENDOCRINOL. 4:329-336(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;

```

RX MEDLINE; 91141487.
RA MOHN K.L., MELEY A.E., TEWARI D.S., LAZ T.M., TAUB R.A.;
RL MOL. CELL. BIOL. 11:1393-1401(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93149132.
RA OOI G.T., TSENG L.Y.H., TRAN M.O., RECHLER M.M.;
RL MOL. ENDOCRINOL. 6:2219-2228(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX MEDLINE; 94250701.
RA LACSON R.G., OEHLER D., YANG E., GOSWAMI R., UNTERMAN T.G.;
RL BIOCHIM. BIOPHYS. ACTA 1218:95-98(1994).
RN [5]
RP SEQUENCE OF 26-59.
RX MEDLINE; 90322923.
RA UNTERMAN T.G., OEHLER D.T., GOTWAY M.B., MORRIS P.W.;
RL ENDOCRINOLOGY 127:789-797(1990).
CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
CC -!- BINDS EQUALLY WELL IGF-I AND IGF-II.
CC -!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
DR EMBL; M58634; G204733; -.
DR EMBL; M89791; G204737; -.
DR EMBL; L22979; G1098473; -.
DR PIR; A36082; A36082.
DR PIR; A37398; A37398.
DR PIR; A39683; A39683.
DR PROSITE; PS00222; IGF BINDING.
DR PROSITE; PS00484; THYROGLOBULIN_1.
KW GROWTH FACTOR BINDING; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 272
FT INSULIN-LIKE GROWTH FACTOR BINDING-
FT PROTEIN 1.
FT DOMAIN 215 264
FT SITE 259 261
FT CELL ATTACHMENT SITE.
FT R -> A (IN REF. 1 AND 4).
FT CONFLICT 79 79
FT AA -> PP (IN REF. 1).
FT CONFLICT 111 112
FT A -> R (IN REF. 3).
FT CONFLICT 201 201
FT H -> N (IN REF. 1).
FT CONFLICT 265 265
SQ SEQUENCE 272 AA; 29684 MW; 59EDF790 CRC32;
Query Match 4.2%; Score 113; DB 5; Length 272;
Best Local Similarity 42.9%; Pred.No. 3.32e-05;
Matches 21; Conservative 4; Mismatches 18; Indels 6; Gaps 6;
Db 43 cpvpaas-cpe-isrp-agcgccptcalplgaacgvatarc-aq-glsc 86
QY 35 CPDEFAPRCFAGVSLVLDGCGCCRCVCAKQLGELC-TERDPCDPKHLFC 82

```

Search completed: Wed Sep 17 09:33:10 1997
Job time : 55 secs.

maryh@stic

4382-1

NeWSprinter20

Wed Sep 17 11:20:33 1997

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5

MPSEARCH

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MParch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 17 09:30:25 1997; MasPar time 7.73 seconds
Tabular output not generated. 516.210 Million cell updates/sec

Title: >US-08-386-680-2
Description: (1-349) from 5585270.pap
Perfect Score: 2713
Sequence: 1 MTAASMGFVRVAFVLLALC.....PGNDIFESLYRKMYGDMA 349

Scoring table: PAM 150
Gap 11

Searched: 96640 seqs, 11439865 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq27
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20

Statistics: Mean 32.775; Variance 141.735; scale 0.231

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	2713	100.0	349	20	Connective tissue gro	3.93e-253
2	2713	100.0	349	16	Connective tissue gro	3.93e-253
3	2521	92.9	348	5	Beta-IG-M2.	4.65e-234
4	1460	53.8	351	6	Chicken nov protein.	4.62e-129
5	1113	41.0	379	5	Beta-IG-M1.	5.50e-95
6	1002	36.9	375	16	Connective tissue gro	3.88e-84
7	644	23.7	205	6	Homologous to chicken	1.93e-49

8	579	21.3	72	6	R31610	Fragment XXVI with bo	3.25e-43
9	430	15.8	84	6	R31602	Chicken nov protein f	4.30e-29
10	412	15.2	71	6	R31609	Encoded by chicken no	2.09e-27
11	405	14.9	84	6	R31603	Polypeptide X homolog	9.43e-27
12	395	14.6	124	10	R46078	CYR61 like protein.	8.10e-26
13	314	11.6	72	6	R31605	Homologous to chicken	2.49e-18
14	312	11.5	70	6	R31604	Chicken nov protein f	3.79e-18
15	278	10.2	75	6	R31601	Chicken nov protein f	4.66e-15
16	268	9.9	76	6	R31600	Chicken nov protein f	3.72e-14
17	197	5.8	22	6	R31612	Fragment XXXI homolog	1.68e-04
18	138	5.1	22	6	R31611	Fragment XXX encoded	5.99e-03
19	125	4.6	271	5	R26994	Rat IGFBP-5.	6.48e-02
20	123	4.5	184	19	R98994	Vascular IBP-like gro	9.29e-02
21	123	4.5	258	4	R22253	Sequence of insulin-1	9.29e-02
22	123	4.5	258	4	R21688	Sequence of insulin-1	9.29e-02
23	121	4.5	455	19	W00231	Drosophila morphogeni	1.33e-01
24	121	4.5	455	19	R47261	Pre-pro 60A.	1.33e-01
25	121	4.5	455	10	R60966	Drosophila 60A morpho	1.33e-01
26	121	4.5	455	10	R57981	Drosophila 60A morpho	1.33e-01
27	121	4.5	455	7	R33413	Drosophila 60A morpho	1.33e-01
28	121	4.5	455	7	R33933	Drosophila 60A morpho	1.33e-01
29	121	4.5	455	9	R47298	Morphogen 60A full le	1.33e-01
30	121	4.5	455	10	R50208	Drosophila 60A morpho	1.33e-01
31	121	4.5	455	10	R60962	Drosophila 60A morpho	1.33e-01
32	121	4.5	455	9	R46745	Drosophila sp. 60(A)	1.33e-01
33	117	4.3	282	13	R79101	Prostaglandin 12 (PGI	2.72e-01
34	117	4.3	282	13	R79102	Prostaglandin 12 (PGI	2.72e-01
35	113	4.2	193	3	P60463	Sequence of C-terminu	5.50e-01
36	113	4.2	2813	3	P60462	Sequence of human von	5.50e-01
37	112	4.1	2813	3	P60053	Sequence of von Wille	6.56e-01
38	109	4.0	1218	19	W05833	Human Serrate-1 (HSL)	1.11e+00
39	105	3.9	272	19	R95329	Insulin-like growth f	2.21e+00
40	105	3.9	272	5	R26995	Human IGFBP-5.	2.21e+00
41	105	3.9	272	5	R25700	IGFBP6.	2.21e+00
42	105	3.9	272	10	R55084	Human insulin-like gr	2.18e+00
43	106	3.9	807	8	R42421	F-spondin (FP5-9).	1.86e+00
44	103	3.8	328	1	R04908	EcoRI-EcoRI fragment	3.11e+00
45	103	3.8	532	17	R83016	Recombinant papilloma	3.11e+00

ALIGNMENTS

RESULT	1
ID	W11302 standard; Protein; 349 AA.
AC	W11302;
DT	18-MAR-1997 (first entry)
DE	Connective tissue growth factor.
KW	Connective tissue growth factor; CTGF; human; connective tissue cell;
KW	proliferative disease; platelet-derived growth factor; PDGF; development;
KW	tissue growth; repair; umbilical vein endothelial cell; HUVE cell;
KW	antibody; wound healing; cancer; fibrotic disease; atherosclerosis;
KW	inhibitor; protease degradation; growth factor; therapy.
OS	Homo sapiens.
PN	US5585270-A.
PD	17-DEC-1996.
PF	30-AUG-1991; 752427.
PR	30-AUG-1991; US-752427.
PR	14-DEC-1993; US-167628.
PR	10-FEB-1995; US-386680.
PA	(UYSF-) UNIV SOUTH FLORIDA.

PI Bradham DM, Grotendorst GR;
DR WPI; 97-051180/05.
DR N-PSDB; T51234.
PT New nucleic acid encoding connective tissue growth factor - useful
PT for accelerating wound healing, also for diagnosis and treatment of
PT proliferative disease
PS Claim 9; Column 15-18; 11pp; English.
CC This sequence represents the human connective tissue growth factor
CC (CTGF). CTGF is related immunologically and biologically to
CC platelet-derived growth factor (PDGF), but is encoded by an unrelated
CC gene. CTGF is thought to play a significant role in the normal
CC development, growth, and repair of human tissue, similarly to PDGF. The
CC cDNA encoding this sequence was isolated by screening a cDNA library from
CC human umbilical vein endothelial (HUVE) cells with anti-PDGF antibodies.
CC CTGF can be used to accelerate wound healing. Also, elevated levels of
CC CTGF may be diagnostic of proliferative diseases involving outgrowth of
CC connective tissue cells, such as cancer, fibrotic disease and
CC atherosclerosis. All of these diseases can be treated with reagents
CC reactive with CTGF, such as antibodies (which can also serve as assay
CC reagents). Antisense nucleic acids, and ribozymes could also be used to
CC inhibit CTGF production. The advantage with using CTGF is that it is
CC more stable, and less susceptible to protease degradation than PDGF, and
CC other growth factors involved in wound healing. This is believed to be
CC due to the high Cys content.
SQ Sequence 349 AA;

Query Match 100.0%; Score 2713; DB 20; Length 349;
Best Local Similarity 100.0%; Pred. No. 3,93e-253;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mtaasgprvrvafvllalcsrpavqncsgpcrcpdepaprcpagvalvldgcgcrcvc 60
QY 1 MTAASGPRVRFVAFVLLALCSRPVAVGQNCSPCRCPDEPAPRCPAGVSLVLDGCGCCRCVC 60
Db 61 akqlgelcterdpchphkgldfcdfgapanrkigvctakdgapciffgtvyrsgesfgasc 120
QY 61 AKQLGELCTERDPCDPHKGILFCDFGSPANRKIGVCTAKDGAFCIFGTVYRSGESFQSSC 120
Db 121 kyqctclldgavcmplcsmdivrlpsdpcfprrvklpgkccceewcdepkdtvvgpala 180
QY 121 KYQCTCLDGAVGCMPLCSMDVRLPSDPCFFPRRVKLPKCCCEEWCDPEKDTVVGPALA 180
Db 181 ayrledtfgpdpmtirancvlqtetwsacsktcgmgiatrvtnndnascrlekqslcmvr 240
QY 181 AYRLEDTFGPDPTMIRANCLVQTETWSACSKTCGMGIATRVTNNDNASCRLEKQSLCMVR 240
Db 241 pceadleenikkkgkciirtpkiskpikfelsgctamkyrakfcgvtcdgrcctphrttt 300
QY 241 PCEADLEENIKKGKCIIRTPKISKPIKFEISGCTSMKYRAKFCGVTCDGRCTPHRTTT 300
Db 301 lpvefkcpdgevmmkmmfiktcaachyncpgndndifesylyrkmvgdma 349
QY 301 LPVEFKCPDGEVMMKMMFIKTCACHYNCPGNDNIFESLYYRKMVGDMA 349

RESULT 2
ID R79964 standard; Protein; 349 AA.
AC R79964;
DT 12-JUN-1996 (first entry)
DE Connective tissue growth factor.

KW Connective tissue growth factor; CTGF; wound healing; vulnary;
KW cell proliferation; cancer; fibrosis; atherosclerosis; diagnosis
OS Homo sapiens.

FH Key Location/Qualifiers
FT Modified site 28
FT /label= N-glycosylation_site
FT Modified site 225
FT /label= N-glycosylation_site
PN US5408040-A.
PD 18-APR-1995; 752427.
PF 30-AUG-1991; 752427.
PR 30-AUG-1991; US-752427.
PR 14-DEC-1993; US-167628.
PA (USF-) UNIV SOUTH FLORIDA.
PI Bradham DM, Grotendorst GR;
DR WPI; 95-161147/21.
DR N-PSDB; T04226.

PT New connective tissue growth factor - used to develop prods. for
PT wound healing and for diagnosis and therapy of cell proliferative
PT disorders.
PS Claim 1; Column 19-20; 12pp; English.
CC Novel human connective tissue growth factor (CTGF) (R79964)
CC is related immunologically and biologically to platelet-derived
CC growth factor (PDGF), but is the product of a distinct gene.
CC CTGF is mitogenic and also a chemotactic agent for cells. It is
CC produced by endothelial and fibroblastic cells, and probably acts
CC as a growth factor in wound healing. Recombinant CTGF can be obtd.
CC by expression of cDNA clone DB60R32 (T04226) in transformed host
CC cells. It is used to accelerate wound healing, and to raise
CC antibodies useful in detecting disorders associated with overgrowth
CC of cells, such as cancer, fibrotic diseases and atherosclerosis.
SQ Sequence 349 AA;

Query Match 100.0%; Score 2713; DB 16; Length 349;
Best Local Similarity 100.0%; Pred. No. 3.93e-253;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mtaasgprvrvafvllalcsrpavqncsgpcrcpdepaprcpagvalvldgcgcrcvc 60
QY 1 MTAASGPRVRFVAFVLLALCSRPVAVGQNCSPCRCPDEPAPRCPAGVSLVLDGCGCCRCVC 60
Db 61 akqlgelcterdpchphkgldfcdfgapanrkigvctakdgapciffgtvyrsgesfgasc 120
QY 61 AKQLGELCTERDPCDPHKGILFCDFGSPANRKIGVCTAKDGAFCIFGTVYRSGESFQSSC 120
Db 121 kyqctclldgavcmplcsmdivrlpsdpcfprrvklpgkccceewcdepkdtvvgpala 180
QY 121 KYQCTCLDGAVGCMPLCSMDVRLPSDPCFFPRRVKLPKCCCEEWCDPEKDTVVGPALA 180
Db 181 ayrledtfgpdpmtirancvlqtetwsacsktcgmgiatrvtnndnascrlekqslcmvr 240
QY 181 AYRLEDTFGPDPTMIRANCLVQTETWSACSKTCGMGIATRVTNNDNASCRLEKQSLCMVR 240
Db 241 pceadleenikkkgkciirtpkiskpikfelsgctamkyrakfcgvtcdgrcctphrttt 300
QY 241 PCEADLEENIKKGKCIIRTPKISKPIKFEISGCTSMKYRAKFCGVTCDGRCTPHRTTT 300
Db 301 lpvefkcpdgevmmkmmfiktcaachyncpgndndifesylyrkmvgdma 349
QY 301 LPVEFKCPDGEVMMKMMFIKTCACHYNCPGNDNIFESLYYRKMVGDMA 349

QY 2 TAASMG-PVRVAVFVLLALCSRAVQNCSGPC--RCDEPAPRCFAGVSLVLDGCGCR 58
Db 62 vcarqgescp1lpedegglycdrgpedgggag1cmlegdncvfdgmryrnetqcp 121
QY 59 VCAKQELGELCTERDPCDPKHGLFCDFGSPANRKGIVCTAKDAPCIFGGTYRSGESFQS 118
Db 122 skyqctcdrgd1gclprnclgl1lpgdpcprfk1evpgcecekwwcd-prdevllggf 180
QY 119 SKYQCTCLDGAAGVCMPLCSMDVRLSPDCPPFRVRKLPKGCCCEWVDEPKDQTVVGP- 177
Db 181 amaayqeatlgldvdsancieqtetwscakscgmfgstrvtrnnrqcmvktqlc 240
QY 178 ALAAYLEDTFGDPDMIRANCLVQTEWSACSKTCGMGISTRVNDNASCRLEKQSRLC 237
Db 241 mmrpeene-epedkkgkciqtckmkavrfeykntavqtkpkygcglndgrcctphn 299
QY 238 MVRPCADLEENIKKGGKCIPTPKIKPIKELSGCTSMKTYRAKFCGVCIDGCCIPHR 297
Db 300 ttiqvefcpggkflkppmlntevchgnpcqennaffqpl 342
QY 298 TTTLPVEFKPCDGEVKKNNMFIKTCACHYNCP-GDNDFESL 339

RESULT 5

ID R25565 standard; Protein; 379 AA.
AC R25565;
DT 18-JAN-1993 (first entry)
DE Beta-IG-M1.
KW Transforming growth factor beta; induced; CEF-10; v-src; chicken;
KW embryo; fibroblasts; TGF-beta.
OS Mus musculus.
PN EP-495674-A.
PD 22-JUL-1992.
PF 17-JAN-1992; 300429.
PR 18-JAN-1991; US-642991.
PR 10-JAN-1992; US-816270.
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Brunner AM, Chinn J, Neubauer MG, Purchio AF;
DR WPI; 92-243508/30.
DR N-PSDB; Q26421.
PT TGF-beta induced gene family - encodes proteins involved in
PT growth and differentiation effects of TGF-beta-1
PS Claim 2; Fig 1; 35pp; English.
CC The protein sequence was deduced from the DNA sequence obt'd. by
CC screening a cDNA library made from AKR-2B mouse cells induced with
CC TGF-beta1 and cyclohexamide with two probes from untreated AKR-2B
CC mRNA and AKR-2B mRNA from cells treated with cyclohexamide and TGF-
CC beta1. The proteins encoded by hybridising colonies (beta-IG-M1 and
CC beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1.
CC Beta-IG-M1 displays 80 percent homology to the CEF-10 protein
CC induced by v-src in chicken embryo fibroblasts and is identical
CC to the protein encoded by cyr61, an immediate early response gene
CC induced in quiescent BALB 3T3 cells by serum treatment. Residues
CC 49-56 of beta-IG-M1 conform to the GCGCCXXC motif reported in the
CC amino half of insulin-like growth factor (IGF) binding proteins.
CC The C-terminal Cys rich region of beta-IG-M1, -M2 and CEF-10 contain
CC an amino acid sequence with strong homology to a motif found near the
CC C-terminal of the malarial circumsporozoite (CS) protein, which is
CC highly conserved among all species of malarial parasites sequenced

CC to date (designated region II). This motif is also found in
CC other proteins which have cell adhesive properties that mediate
CC cell-cell and cell-extracellular matrix interactions, such as
CC properdin, thrombospondin, and tRAP. The proteins encoded by
CC TGF-beta induced genes are likely to be involved in mediation of
CC the biological effects of TGF-beta relating to cell growth and
CC differentiation. See also R25566.
SQ Sequence 379 AA;

Query Match 41.0%; Score 1113; DB 5; Length 379;
Best Local Similarity 48.3%; Pred. No. 5.50e-95;
Matches 180; Conservative 79; Mismatches 76; Indels 38; Gaps 20;

Db 1 messtfrt1ava-vtllhl-trials-tcpaacheple-apkcapgvglvrdgcgckvc 56
QY 1 MTAASMGFVRVAVFVLLALCSRAVQNCSGPCRCDEPAPRCFAGVSLVLDGCGCRVC 60
Db 57 akqinedcsktqpcdhtkglecnfgaastalkgicraqeegrpceynsrlngesafqn 116
QY 61 AKQELGELCTERDPCDPKHGLFCDFGSPANRKGIVCTAK-DGAPCIFGGTYRSGESFQS 119
Db 117 khqctcidgavgcip1cpqelalplnlgpnp1rvkvsgqcccewcdedskldldgd 176
QY 120 KYQCTCLDGAAGVCMPLCSMDVRLSPDCPPFRVRKLPKGCCCEWVDE-P-KD----QT 173
Db 177 dl1gldaeveltrnnel1aigksal1krlpvf1gt1p1v1fn1l1h1g1k1c1v1q1t1t1s1w1s1q 236
QY 174 -VVG--PA---LA-----AY-R---LE--DTFGPDP-TM---IRAN---CIVQTTEWSA 208
Db 237 cskscgtg1strvndnpeclrvktricevpcgppvysalkkkgkcksktkkspvrf 296
QY 209 CSKTCGMGISTRVNDNASCRLEKQSRLCVRFCEADLEENIKKGGKCIPTPKISKPIRF 268
Db 297 tyagcsvskyrpkygcsgvdgrcctplqtrtvkmrfrcecdgemfaknmvliqsccknyn 356
QY 269 ELSGCTSMKTYRAKFCGVCIDGCCIPHRTTTLTPVEFKPCDGEVKKNNMFIKTCACHYN 328
Db 357 cphpneaf-rly 368
QY 329 CPGNDI-FESLY 340

RESULT 6

ID R90919 standard; Protein; 375 AA.
AC R90919;
DT 25-JUN-1996 (first entry)
DE Connective tissue growth factor-2.
KW CTGF-2; connective tissue growth factor-2.
KW cartilaginous growth; skeletal; embryo; cell growth; morphogenesis;
KW insulin-like growth factor; fibroblast growth factor; Cry61.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..24
FT /label= signal_peptide
FT Protein 25..375
FT /label= mature_protein
PN W09601896-A.
PD 25-JAN-1996.
PF 12-JUL-1994; U07736.
PR 12-JUL-1994; W0-U07736.

CC probe. Nucleotide sequences of the invention include those which
 CC code for a protein having at least 60% homology with the protein
 CC fragment XXV (R31609) deduced from the 3rd. exon of the chicken nov
 CC gene. In particular, the sequences of the invention encode a
 CC protein having the sequence of Fragment XXVI (R31610). (Sequence
 CC XXVI is described as a nucleotide sequence in the claims but is
 CC correctly described as an amino acid sequence in the disclosure).
 SQ Sequence 72 AA;

Query Match 21.3%; Score 579; DB 6; Length 72;
 Best Local Similarity 97.2%; Pred. No. 3.25e-43;
 Matches 70; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 dgapcifgtvyrsgesfqsckytctldgavgcmlcsmdvrlpsdpcpfrvklpg 60
 Qy 99 DGAPCIFGTVYRSGESFQSCKYQCTCLDGA VGCMPLCSMDVRLPSDPCPFRVKLP 158
 Db 61 mceewwcvdepr 72
 Qy 159 KCCEEWVCDEPK 170

RESULT 9

ID R31602 standard; Protein; 84 AA.
 AC R31602;
 DT 24-MAY-1993 (first entry)
 DE Chicken nov protein fragment IX.
 KW avian nephroblastoma; avian myeloblastoma virus;
 KW stringent hybridisation.
 OS Gallus domesticus.
 PN W09300430-A.
 PD 07-JAN-1993.
 PF 25-JUN-1992; F00589.
 PR 25-JUN-1991; FR-007807.
 PA (CNRS) CENT NAT RECH SCI.
 PI Martinerie C. Perbal B;
 DR WPI; 93-036377/04.
 DR N-PSDB; Q36035.
 PT Nucleotide sequences hybridising to regions of chicken nov gene -
 PT useful as probes for detecting complementary sequences to
 PT evaluate development and/or differentiation of tumours
 PS Claim 8; Page 31; 67pp; French.
 CC The chicken nov gene is stimulated in avian nephroblastoma induced
 CC by avian myeloblastoma virus but not in normal adult kidney. A
 CC 1975bp cDNA sequence (Q36035) was isolated from a gene bank prepared
 CC from chicken embryonic fibroblasts screened with a tumour-derived
 CC probe. Nucleotide sequences of the invention include those which
 CC code for a protein having at least 60% homology with the protein
 CC fragment VIII (Q36035) is derived from the 3rd. exon of the nov gene;
 CC nucleotide sequences which hybridise to Fragment VIII under stringent
 CC conditions (i.e. 50% formamide, 5 x SCC) are claimed. The claimed
 CC sequences preferably encode a protein with at least 70% homology to
 CC amino acid sequence IX (R31602) which is encoded by Fragment VIII.
 SQ Sequence 84 AA;

Query Match 15.8%; Score 430; DB 6; Length 84;
 Best Local Similarity 57.7%; Pred. No. 4.30e-29;
 Matches 45; Conservative 19; Mismatches 13; Indels 1; Gaps 1;

Db 3 egdncvfdgmlyrnetfqpckvqctctdggcgcprcnlglllpdpdpfrkivpg 62
 Qy 99 DGAPCIFGTVYRSGESFQSCKYQCTCLDGA VGCMPLCSMDVRLPSDPCPFRVKLP 158

Db 63 eccekwvcd-prdevllg 79
 Qy 159 KCCEEWVCDEPKQIVVG 176

RESULT 10

ID R31609 standard; Protein; 71 AA.
 AC R31609;
 DT 24-MAY-1993 (first entry)
 DE Encoded by chicken nov gene exon 3 fragment.
 KW avian nephroblastoma; avian myeloblastoma virus;
 KW stringent hybridisation.
 OS Gallus domesticus.
 PN W09300430-A.
 PD 07-JAN-1993.
 PF 25-JUN-1992; F00589.
 PR 25-JUN-1991; FR-007807.
 PA (CNRS) CENT NAT RECH SCI.
 PI Martinerie C. Perbal B;
 DR WPI; 93-036377/04.
 PT Nucleotide sequences hybridising to regions of chicken nov gene -
 PT useful as probes for detecting complementary sequences to
 PT evaluate development and/or differentiation of tumours
 PS Claim 25; Page 40; 67pp; French.
 CC The chicken nov gene is stimulated in avian nephroblastoma induced
 CC by avian myeloblastoma virus but not in normal adult kidney. A
 CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
 CC from chicken embryonic fibroblasts screened with a tumour-derived
 CC probe. Nucleotide sequences of the invention include those which
 CC code for a protein having at least 60% homology with the protein
 CC fragment XXV (R31609) deduced from the 3rd. exon of the chicken nov
 CC gene. (Sequence XXV is described as a nucleotide sequence in the
 CC claims but is correctly described as an amino acid sequence in the
 SQ Sequence 71 AA;

Query Match 15.2%; Score 412; DB 6; Length 71;
 Best Local Similarity 59.7%; Pred. No. 2.09e-27;
 Matches 43; Conservative 16; Mismatches 12; Indels 1; Gaps 1;

Db 1 egdncvfdgmlyrnetfqpckvqctctdggcgcprcnlglllpdpdpfrkivpg 60
 Qy 99 DGAPCIFGTVYRSGESFQSCKYQCTCLDGA VGCMPLCSMDVRLPSDPCPFRVKLP 158
 Db 61 eccekwvcd-pr 71
 Qy 159 KCCEEWVCDEPK 170

RESULT 11

ID R31603 standard; Protein; 84 AA.
 AC R31603;
 DT 24-MAY-1993 (first entry)
 DE Polypeptide X homologous to chicken nov protein fragment.
 KW avian nephroblastoma; avian myeloblastoma virus;
 KW stringent hybridisation.
 OS Homo sapiens.
 PN W09300430-A.
 PD 07-JAN-1993.

<p> ID AC R31604; 24-MAY-1993 DT (first entry) DE Chicken nov protein fragment XIV. KW avian myeloblastoma virus; OS avian nephroblastoma; avian myeloblastoma virus; PN stringent hybridisation. Gallus domesticus. W09300430-A. 07-JAN-1993. 25-JUN-1992; F00589. PR 25-JUN-1991; FR-007807. CNR(S) CENT NAT RECH SCI. PI Martinerie C, Perbal B; WPI; 93-036377/04. DR N-PSDB; Q36038. PDR Nucleotide sequences hybridising to regions of chicken nov gene - PR useful as probes for detecting complementary sequences to PPT evaluate development and/or differentiation of tumours PPT Claim 13; Page 34; 67pp; French. PFS The chicken nov gene is stimulated in avian nephroblastoma induced CCC by avian myeloblastoma virus but not in normal adult kidney. A CCC 1975bp cDNA sequence was isolated from a gene bank prepared from CCC chicken embryonic fibroblasts screened with a tumour-derived probe. CCC Fragment XII (Q36038) is derived from the 4th. exon of the nov gene; CCC nucleotide sequences which hybridise to Fragment XIII under stringent CCC conditions (i.e. 50% formamide, 5 x SCC) are claimed. The claimed CCC sequences preferably encode a protein with at least 86% homology to CCC amino acid sequence XIV (R31604) which is encoded by Fragment XIII. SQ Sequence 70 AA; </p>	<p> R31604 standard; Protein; 70 AA. </p>
--	---

Query Match	11.5%	Score 312;	DB 6;	Length 70;
Best Local Similarity	64.6%	Pred. No. 3.79e-18;		
Matches	42;	Conservative	9;	Mismatches 14;
			Indels	0;
			Gaps	0;

```

Db      1  ayrqeatlglidvdsansancieqttewsacsckgmgfstrvnrnngacemvkqtrllcmrr 60
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Oy     181 AYRLDTFGDPNIRANCLVQTTEWSACSKTCGMGISTRTVNDNASCELEKOSRLCQWR 240

```

Db	61	pcene	65
		:	
Qv	241	PCEAD	245

RESULT	15	
ID	R31601	standard; Protein; 75 AA.
AC	R31601;	
DE	24-MAY-1993	(first entry)
DE	Chicken nov protein fragment V.	
DE	avian nephroblastoma; avian myeloblastoma virus;	
KW	stringent hybridisation.	
KW	Gallus domesticus.	
OS	Gallus domesticus.	
PN	WO9300430-A.	
PD	07-JAN-1993.	
PF	25-JUN-1992;	FOO589.
PPR	25-JUN-1991;	FR-007807.
PA	(CNRS) CENT NAT RECH SCI.	
PI	Martinerie C,	
PI	Perbal B;	
DR	WP; 93-Q36377/04.	
PT	Nucleotide sequences hybridising to regions of chicken nov gene	
PT	useful as probes for detecting complementary sequences to	
PT	nov	

PT evaluate development and/or differentiation of tumours
PS Claim 5; Page 28; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Nucleotide sequences which hybridise to Q36031 or specified
CC sub-fragments of it, under stringent conditions (i.e. 50% formamide,
CC 5 x SSC), are claimed. The claimed sequences preferably encode a
CC protein with amino acid sequence V (R31601).
SQ Sequence 75 AA;

Query Match 10.2%; Score 278; DB 6; Length 75;
Best Local Similarity 48.0%; Pred. No. 4,66e-15;
Matches 36; Conservative 14; Microstates 22.

2 aatgrrppqpcgrcpatp-ptcapgvrvldgcscclvcarggescsdlepcdessgly 60
| | : | | : | | : | | : | | : | | : | | : | | : | | :
24 AVGQNCSGFPC--RCPDEAPRCPAGVSLVDGGCCRCVCAKQIGELCTERDPCDPHKGLF 81

Db	61	cdrsadpsnqtgict	75
		:: :: ::	
OY	82	CDFGSPANRKIGVCT	96

Search completed: Wed Sep 17 09:31:56 1997
Job time : 91 secs.

maryh@stic

4383-1

NeWSprinter20

Wed Sep 17 11:21:24 1997

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5

MPARCH_PP protein - protein database search, using Smith-Waterman algorithm

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPARCH_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 17 09:28:51 1997; MasPar time 13.03 Seconds
Tabular output not generated.

Title: >US-08-386-680-2
Description: (1-349) from 5565270.pep
Perfect Score: 2713
Sequence: 1 MTAASMGFVRVAFVLLALC.....PGDNDIFSLYRKMGDMA 349

Scoring table: PAM 150
Gap 11

Searched: 91006 seqs, 28888923 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir51
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
13:unann9 14:unann10 15:unann16:unrev

Statistics: Mean 42.720; Variance 82.202; scale 0.520

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2713	100.0	349	13	A40551	connective tissue gr 0.00e+00
2	2521	92.9	348	14	A40578	beta IG-M2 protein p 0.00e+00
3	2514	92.7	348	14	A53228	fisp-12 protein prec 0.00e+00
4	1460	53.8	351	13	S20078	NOV protein - chicke 3.42e-284
5	1415	52.2	357	13	I38069	gene novH protein - 2.53e-274
6	1152	42.5	375	13	A41428	CEF-10 protein precu 7.43e-217
7	1113	41.0	379	14	A55669	gene CYR61 protein p 2.27e-208
8	133	4.9	254	14	JC1464	insulin-like growth 1.32e-06

9	133	4.9	254	14	I48599	insulin-like growth 1.32e-06
10	128	4.7	271	13	JC4584	insulin-like growth 8.02e-06
11	125	4.6	271	14	JC1463	insulin-like growth 2.33e-05
12	126	4.6	835	13	JP0076	nel protein - chicke 1.63e-05
13	123	4.5	238	13	B37252	insulin-like growth 4.71e-05
14	121	4.5	455	6	A43918	TGF-beta-related pro 9.49e-05
15	116	4.3	254	14	I48603	insulin-like growth 5.32e-04
16	117	4.3	277	13	I52825	gene MAC25 protein - 3.78e-04
17	117	4.3	282	13	S50031	prostacyclin-stimula 3.78e-04
18	118	4.3	1348	12	A43917	probable epidermal c 2.68e-04
19	118	4.3	1348	12	S27812	insulin-like growth 7.48e-04
20	115	4.2	272	6	A36082	von Willebrand facto 1.47e-03
21	113	4.2	2813	3	VWHD	insulin-like growth 4.01e-03
22	110	4.1	263	6	S23009	insulin-like growth 2.88e-03
23	111	4.1	271	14	I48604	insulin-like growth 2.88e-03
24	111	4.1	304	6	A33274	insulin-like growth 2.88e-03
25	111	4.1	305	14	I48601	insulin-like growth 2.88e-03
26	111	4.1	305	6	JN0508	mucin 2 precursor, i 4.01e-03
27	110	4.1	3020	13	A43932	insulin-like growth 5.59e-03
28	109	4.0	124	6	S25113	insulin-like growth 5.59e-03
29	109	4.0	272	14	I48600	insulin-like growth 5.59e-03
30	106	3.9	258	14	A45403	insulin-like growth 1.49e-02
31	105	3.9	272	13	A53748	insulin-like growth 2.07e-02
32	106	3.9	807	14	A38152	F-spondin - rat 1.49e-02
33	102	3.8	259	2	IOHD1	insulin-like growth 5.40e-02
34	102	3.8	291	2	JN0064	insulin-like growth 5.40e-02
35	104	3.8	310	6	A60967	insulin-like growth 2.85e-02
36	103	3.8	328	2	A41927	insulin-like growth 3.93e-02
37	103	3.8	500	4	P1WL6	I1 protein - human p 3.93e-02
38	102	3.8	798	7	A40526	integrin beta-7 chai 5.40e-02
39	104	3.8	3133	16	S70920	lectin - silkworm 2.85e-02
40	104	3.8	3133	16	S52093	circumsporozoite pro 7.42e-02
41	101	3.7	388	7	A39756	prepro osteogenic pr 7.42e-02
42	101	3.7	402	6	A45056	circumsporozoite pro 7.42e-02
43	101	3.7	405	7	S05428	circumsporozoite pro 7.42e-02
44	101	3.7	412	3	OZ20AF	circumsporozoite pro 7.42e-02
45	101	3.7	424	12	A54533	circumsporozoite pro 7.42e-02

ALIGNMENTS

RESULT ENTRY	#type complete
connective tissue growth factor - human	A40551
#formal_name Homo sapiens #common name man	connective tissue growth factor
17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change	25-Aug-1995
ACCESSIONS	A40551; S44205
REFERENCE	A40551
#authors	Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotendorst, G.R.
#journal	J. Cell Biol. (1991) 114:1285-1294
#title	Connective tissue growth factor: a cysteine-rich mitogen secreted by human vascular endothelial cells is related to the SRC-induced immediate early gene product CEF-10.
#cross-references	WMD:91373462
#accession	A40551
#molecule_type	mRNA
#residues	1-349
REFERENCE	S44205
#authors	Oemar, B.S.; Werner, A.; Yang, Z.; Garnier, J.M.; Gentz, R.

Luescher, T.F.
#submission submitted to the EMBL Data Library, April 1994
#description Differential cloning and expression of human connective
tissue growth factor.
#accession S44205
#status preliminary
#molecule_type mRNA
#residues 1-349 #label OEM
#cross-references EMBL:X78947
#length 349 #molecular-weight 38069 #checksum 8930

SUMMARY

Query Match 100.0%; Score 2713; DB 13; Length 349;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 1 mtaasmgprvrvafvllalcsrpavqncsgpcrcpdepaprcpagvalvldgcgcrcvc 60
QY 1 MTAASMGPRVAVFVLLALCSRPVAVGQNSGPGRCDEPAPRCPCPAGVSLVLDGCGCCRCVC 60

Db 61 akqlgelcterdpcdphkglfcdgspnarkigvctakdgapciffgtgvyrsgefsqsc 120
QY 61 AKQLGELCTERDPCDPHKGLFCDFGSPANRKGVCCTAKDGAPCIIFFGTGVRSGESFQSSC 120

Db 121 kyqctcldgavgcmlpcsdvrlpsdpcfprrvklpgkceewvcdpdkdgtvvvgpala 180
QY 121 KYQCTCLDGAVGCMPLCSMDVRLPSDPCFPRRVKLPKGKCEEWVCDPEPKDQTVVGPALA 180

Db 181 avrledtfgpdpmtiranclvqtewascktcgmgiistrvndnascrlekqsrclmvr 240
QY 181 AYRLEDTFGDPDMTIRANCLVQTEWSACSKTCGMGISTRTVNDNASCRLEKQSRCLMVR 240

Db 241 pceadleenikgkckirtpkiskplkfelsgctsmkyrakfcgvctdgrcctphrttt 300
QY 241 PCEADLEENIKGKCKIRTPKISKPIKFLSGCTSMKYRAKFCGVCTDGRCCTPHRTTT 300

Db 301 lpvefkcpdgevmkmmfiktacachncpgdndifesliyrkmygdma 349
QY 301 LPVEFKCPDGEVMKMMFIKTACHNCPGDNDIFESLIYRKMVGMDMA 349
```

RESULT 2
ENTRY beta IG-M2 protein precursor - mouse
TITLE
ORGANISM
DATE 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 30-Sep-1993
ACCESSIONS A40578
REFERENCE A40578
#authors Brunner, A.; Chinn, J.; Neubauer, M.; Purchio, A.F.
#journal DNA Cell Biol. (1991) 10:293-300
#title Identification of a gene family regulated by transforming growth factor-beta.
#cross-references M01D:91229699
#accession A40578

SUMMARY
#status preliminary
#molecule_type mRNA
#residues 1-348 #label BRU
#length 348 #molecular-weight 37794 #checksum 9188

Query Match 92.9%; Score 2521; DB 14; Length 348;

Best Local Similarity 91.1%; Pred. No. 0.00e+00;
Matches 318; Conservative 21; Mismatches 9; Indels 1; Gaps 1;

```
Db 1 mlasvagiislaiv-llalctrpatgdcscagcqaacaaahpcpagvalvldgcgcrcvc 59
QY 1 MTAASMGPRVAVFVLLALCSRPVAVGQNSGPGRCDEPAPRCPCPAGVSLVLDGCGCCRCVC 60

Db 60 akqlgelcterdpcdphkglfcdgspnarkigvctakdgapciffgtgvyrsgefsqsc 119
QY 61 AKQLGELCTERDPCDPHKGLFCDFGSPANRKGVCCTAKDGAPCIIFFGTGVRSGESFQSSC 120

Db 120 kyqctcldgavgcmlpcsdvrlpsdpcfprrvklpgkceewvcdpdkdrtavgpala 179
QY 121 KYQCTCLDGAVGCMPLCSMDVRLPSDPCFPRRVKLPKGKCEEWVCDPEPKDQTVVGPALA 180

Db 180 avrledtfgpdpmtiranclvqtewascktcgmgiistrvndntfcrlekqsrclmvr 239
QY 181 AYRLEDTFGDPDMTIRANCLVQTEWSACSKTCGMGISTRTVNDNASCRLEKQSRCLMVR 240

Db 240 pceadleenikgkckirtpkiskplkfelsgctsmkyrakfcgvctdgrcctphrttt 299
QY 241 PCEADLEENIKGKCKIRTPKISKPIKFLSGCTSMKYRAKFCGVCTDGRCCTPHRTTT 300

Db 300 lpvefkcpdgevmkmmfiktacachncpgdndifesliyrkmygdma 348
QY 301 LPVEFKCPDGEVMKMMFIKTACHNCPGDNDIFESLIYRKMVGMDMA 349
```

RESULT 3
ENTRY fisp-12 protein precursor - mouse
TITLE
ORGANISM
DATE 19-May-1994 #sequence_revision 19-May-1994 #text_change 19-May-1994
ACCESSIONS A53228
REFERENCE A53228
#authors Ryaack, R.P.; Macdonald-Bravo, H.; Mattei, M.G.; Bravo, R.
#journal Cell Growth Differ. (1991) 2:225-233
#title Structure, mapping, and expression of fisp-12, a growth factor-inducible gene encoding a secreted cysteine-rich protein.

#accession A53228
#status preliminary
#molecule_type DNA
#residues 1-348 #label RYS
#cross-references GB:M70641

GENETICS
#gene fisp-12
SUMMARY #length 348 #molecular-weight 37793 #checksum 9470

Query Match 92.7%; Score 2514; DB 14; Length 348;
Best Local Similarity 90.8%; Pred. No. 0.00e+00;
Matches 317; Conservative 21; Mismatches 10; Indels 1; Gaps 1;

```
Db 1 mlasvagiislaiv-llalctrpatgdcscagcqaacaaahpcpagvalvldgcgcrcvc 59
QY 1 MTAASMGPRVAVFVLLALCSRPVAVGQNSGPGRCDEPAPRCPCPAGVSLVLDGCGCCRCVC 60

Db 60 akqlgelcterdpcdphkglfcdgspnarkigvctakdgapciffgtgvyrsgefsqsc 119
QY 61 AKQLGELCTERDPCDPHKGLFCDFGSPANRKGVCCTAKDGAPCIIFFGTGVRSGESFQSSC 120
```

Qy 61 AKQGEICTERDPCDPKHGLFCDFGSPANKRIGVCTAKGAPCIFGGTVYRSGESFQSSC 120
 Db 120 kyqctcidgavgevpplcsmdivrpspcdpfprvk1pgkckewcdcdpdkrtavgpala 179
 Qy 121 KYQCTCLDGAAGCPLCSMDVRLSPDCPFPRVRLKFGKCEEWCDKDFDQTVVGPALA 180
 Db 180 ayrltdfpgdptmmrancelvqtwsacaktcmgmistrtvndntfcrlekqslcmvr 239
 Qy 181 AYRLDFTGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVINDNASCRLEKQSLCMVR 240
 Db 240 pceadleenikgkciirtpkiaakpvkfelsgtsvktyrakfgvctdgrctphrttt 299
 Qy 241 PCEADLEENIKGKKCIIRTPKISKP IKFELSGCTSMKTYRAKFCGVTGDRCTPHRTTT 300
 Db 300 lpvefkpcdgeimkmmfiktctachyncpgndndifesiyyrkmvgdma 348
 Qy 301 LPVEFKPCDGEVKKNNMFIKTCACHYNCPGNDNDFESIYYRKMVGDMA 349

RESULT 4
 ENTRY S20078 #type complete
 TITLE NOV protein - chicken
 ORGANISM #formal name Gallus gallus #common name chicken
 DATE 19-feb-1994 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995

ACCESSIONS S20078
 REFERENCE S20078
 #authors Joliot, V.; Martinier, C.; Dambrine, G.; Plassiart, G.; Brisac, M.; Crochet, J.; Perbal, B.
 #journal Mol. Cell. Biol. (1992) 12:10-21
 #title Proviral rearrangements and overexpression of a new cellular gene (nov) in myeloblastosis-associated virus type 1-induced neoplastomas.
 #cross-references MVID:92107157
 #accession S20078
 #status preliminary
 #molecule type mRNA
 #residues 1-351 #label JOL
 #cross-references EMBL:X59284

GENETICS
 #gene NOV
 #note #length 351 #molecular-weight 38268 #checksum 4069

SUMMARY
 Query Match 53.8%; Score 1460; DB 13; Length 351;
 Best Local Similarity 54.2%; Pred. No. 3.42e-284;
 Matches 186; Conservative 69; Mismatches 80; Indels 8; Gaps 7;

Db 3 tgggqgplvllllllrpceavagreaacprpcggrcaep-prcapgvavldgqcgcl 61
 Qy 2 TAASMG-PVRVAVVLLALCSRPVAGQNSGFC--RCPDEPAPRCAGVSLVLDGCGCCR 58
 Db 62 vcarqreacspllpcdeagglvdrpdedggagacmvlegdhcvcfdgmlyrnetfqp 121
 Qy 59 VCAQKQELCTERDPCDPKHGLFCDFGSPANKRIGVCTAKGAPCIFGGTVYRSGESFQS 118
 Db 122 sckyactcdgicgicpncnlgllpgdpdpfprkievepgecekwcdd-prdevllggf 180
 Qy 119 SKCYOCTCLDGAAGCPLCSMDVRLSPDCPFPRVRLKFGKCEEWCDKDFDQTVVGP- 177
 Db 181 amaaayrqeatlgidvdsaaancieqtetwsacskcgmgtftrvnrnqgcmvkgqrlc 240

Qy 178 ALAAVRLDFTGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVINDNASCRLEKQSLC 237
 Db 241 mmpcene-epedkkgkciqtckkamkavrfeykncstvtqkpryogclndgrctphn 299
 Qy 238 MYRPEADLEENIKGKKCIIRTPKISKP IKFELSGCTSMKTYRAKFCGVTGDRCTPHR 297
 Db 300 tktigvefcgkglfkppmmlintcvcgncpqsennaffql 342
 Qy 298 TTTLVPEFKPCDGEVKKNNMFIKTCACHYNCP-GDNDIFESL 339

RESULT 5
 ENTRY I38069 #type complete
 TITLE gene novH protein - human
 ORGANISM #formal name Homo sapiens #common name man
 DATE 17-May-1996 #sequence_revision 17-May-1996 #text_change 17-May-1996

ACCESSIONS I38069
 REFERENCE I38069
 #authors Martinier, C.; Huff, V.; Joubert, I.; Badzioch, M.; Saunders, G.; Strong, L.; Perbal, B.
 #journal Oncogene (1994) 9:2729-2732
 #title Structural analysis of the human nov proto-oncogene and expression in Wilms tumor.
 #cross-references MVID:94336229
 #accession I38069
 #status preliminary; translated from GB/EMBL/DBJ
 #molecule type DNA
 #residues 1-357 #label RES
 #cross-references EMBL:X78351; NID:g587422; CDS_PID:g825696

GENETICS
 #introns 28/3; 104/1; 188/1; 259/3
 #note gene name novH
 #summary #length 357 #molecular-weight 39162 #checksum 1640

Query Match 52.2%; Score 1415; DB 13; Length 357;
 Best Local Similarity 51.8%; Pred. No. 2.53e-274;
 Matches 177; Conservative 68; Mismatches 91; Indels 6; Gaps 5;

Db 18 ltfll-llhllgqvaaatgropqppgpcp-tpcagpvavldgscclvcargesc 75
 Qy 11 VAFVLLALCSRPVAGQNSGFC--RCPDEPAPRCAGVSLVLDGCGCERCAKUGELC 68
 Db 76 sdlepcdeseglycdreadpsndgtictavegdnvfdgvlyrsgkfgpsckfgctcd 135
 Qy 69 TERDPCDPKHGLFCDFGSPANKRIGVCTAKGAPCIFGGTVYRSGESFQSSCKYQCTCLD 12
 Db 136 gqicvprcldvlpepncpaprkvvepgecekwcicgpcdeedalgltlaayrpeatl 195
 Qy 129 GAVGCMPLCSMDVRLSPDCPFPRVRLKFGKCEEWCDKDFDQTVVGPALAAVRLDFT 188
 Db 196 qvevdsasvncieqtetwsacskcgmgtftrvnrnqgcmvkgqrlc 255
 Qy 189 GPDTMIRANCLVQTTEWSACSKTCGMGISTRVINDNASCRLEKQSLCMVRPCADLEE 248
 Db 256 ptdkkgkciirtkkskathlqfknctslhtykrfcgvcsgdgrctphntktiaefqc 315
 Qy 249 NI-KKGGKCIIRTPKISKP IKFELSGCTSMKTYRAKFCGVTGDRCTPHRTTLPVFKC 307

[illegible]

TITLE	gene CYR61 protein precursor - mouse
ORGANISM	#formal_name Mus musculus #common_name house mouse
DATE	28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 18-Oct-1996
ACCESSIONS	A35669; I48319
REFERENCE	A35669
#authors	O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.
#journal	Mol. Cell. Biol. (1990) 10:3569-3577
#title	Expression of cyr61, a growth factor-inducible immediate-early gene.
#cross-references	MDID:90287146
#accession	A35669
#status	preliminary
#molecule_type	mRNA
#residues	1-379 #label OAB
#cross-references	GB:M32490
#note	the authors translated the codon GAT for residue 337
REFERENCE	Gln
I48319	
#authors	Latinkic, B.V.; O'Brien, T.P.; Lau, L.F.
#journal	Nucleic Acids Res. (1991) 19:3261-3267
#title	Promoter function and structure of the growth factor-inducible immediate early gene cyr61.
#cross-references	MDID:91288203
#accession	I48319
#status	translated from GB/EMBL/DBJ
#molecule_type	DNA
#residues	1-379 #label RES
#cross-references	EMBL:X56790; NID:g50632; CDS_PID:g50633
GENETICS	
#gene	CYR61
#introns	21/3; 93/1; 208/1; 279/3
SUMMARY	#length 379 #molecule-weight 41709 #checksum 3726
Query Match	41.0%; Score 1113; DB 14; Length 379;
Best Local Similarity	48.3%; Pred. No. 2.27e-208;
Matches	180; Conservative 79; Mismatches 76; Indels 38; Gaps 3
Db	1 messtfrlava-vtlhl-trials-tcpaachpale-apkcapgvglvrdgcgcckvc 56
QY	1 MTAASMGFVRVAVFVLLALCSRPAVQNGSGCRCPDEPAPCFAGVSLVDGGCCRCV 60
Db	57 akqlnedsktqpdctkglecnfgasstalgkicragsegprceynsriyqngesqpn 116
QY	61 AKQLGELCTERDPCDHPKGLFCDFGSPANRKGIVCTAK-DGAPCIFGTVYRSGESFQSS 119
Db	117 khqctcidgavgcipicpaelplgpcnprlvkvagqceewcdedsikdsldddq 178
QY	120 CKYQCTLDJGAVGCMPLCSMDVRLPSPDFPRVRKLPKGCCEEWCD-E-P-KD-----QT 173
Db	177 dlilgdaaeveltrnnelaighksalrplvfgteprvlfnplbahqdkciavttaweq 236
QY	174 -VVG---FA---LA-----AY-R---LE---DTFGDPP-TM---IRAN---CLVQTTEWSA 208
Db	237 cskscgtqistrvtndnpecirvketricvepcgqpyvsalggkcksktkskspvrf 296
QY	209 CSKTCGMGISTRVNDNASCRLEKQSRCLMWPFPCEADLEENIKKKKCIKTPKISPKIF 268
Db	297 tyagcsavkkyrpkycgscvdgrcetplqtrtvtkmrfcedgemfksnmvmilgckcnyn 356

Qy 269 ELSGCTSMKYTRAFGVCVTDRCCTPHRTITLVEFKCPDGEVKKNNMFIKTCACHYN 328

Db 357 cphneaf-rly 368
| | : | |

Qy 329 CPDNDI-FESLY 340

RESULT 8

ENTRY JCI464 #type complete

TITLE insulin-like growth factor-binding protein 4 precursor - rat

ORGANISM #formal name Rattus norvegicus #common name Norway rat

DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-1995

ACCESSIONS JCI464; A37252; E40403; A61120; B33570; A61118

REFERENCE JCI464

#authors Gao, L.; Ling, N.; Shimasaki, S.

#journal Biochem. Biophys. Res. Commun. (1993) 190:1053-1059

#title Structure of the insulin-like growth factor binding protein-4 gene.

#accession JCI464

#molecule_type DNA

#residues 1-254 #label GAO

#cross-references GB:L08276

REFERENCE A37252

#authors Shimasaki, S.; Uchiyama, F.; Shimonaka, M.; Ling, N.

#journal Mol. Endocrinol. (1990) 4:1451-1458

#title Molecular cloning of the cDNAs encoding a novel insulin-like growth factor-binding protein from rat and human.

#cross-references MUID:91133415

#accession A37252

#molecule_type mRNA

#residues 1-254 #label SHI

REFERENCE A40403

#authors Shimasaki, S.; Shimonaka, M.; Zhang, H.P.; Ling, N.

#journal J. Biol. Chem. (1991) 266:10646-10653

#title Identification of five different insulin-like growth factor binding proteins (IGFBPs) from adult rat serum and molecular cloning of a novel IGFBP-5 in rat and human.

#cross-references MUID:91244847

#accession E40403

#molecule_type protein

#residues 22-26, 'X', 28-29, 'X', 31-37, 'X', 39-40 #label SH2

REFERENCE A61120

#authors Ceda, G.P.; Fielder, P.J.; Henzel, W.J.; Louie, A.; Donovan, S.M.; Hoffman, A.R.; Rosenfeld, R.G.

#journal Endocrinology (1991) 128:2815-2824

#title Differential effects of insulin-like growth factor (IGF)-I and IGF-II on the expression of IGF binding proteins (IGFBPs) in a rat neuroblastoma cell line: isolation and characterization of two forms of IGFBP-4.

#accession A61120

#molecule_type protein

#residues 'X', 23-26, 'X', 28-29, 'X', 31-37, 'X', 39-43, 'X', 45, 'X', 47-48 #label CED

#note both glycosylated and nonglycosylated forms of this protein were observed

REFERENCE A33570

#authors Shimonaka, M.; Schroeder, R.; Shimasaki, S.; Ling, N.

#journal Biochem. Biophys. Res. Commun. (1989) 165:189-195

#title Identification of a novel binding protein for insulin-like

growth factors in adult rat serum.

#cross-references MUID:90073708

#accession B33570

#molecule_type protein

#residues 22-26, 'X', 28-29, 'X', 31-37, 'X', 39-43, 'X', 45-52, 'X', 54, 'XX', 57-58, 'X', 60-61 #label SH3

REFERENCE A61118

#authors Cheung, P.T.; Smith, E.P.; Shimasaki, S.; Ling, N.; Chernausek, S.D.

#journal Endocrinology (1991) 129:1006-1015

#title Characterization of an insulin-like growth factor binding protein (IGFBP-4) produced by the B104 rat neuronal cell line: chemical and biological properties and differential synthesis by sublines.

#accession A61118

#molecule_type protein

#residues 22-26, 'X', 28-29, 'XX', 32-33 #label CHE

GENETICS 117/1; 165/3; 210/3

CLASSIFICATION #superfamily thyroglobulin type I repeat homology

KEYWORDS glycoprotein

FEATURE 1-21

22-254

170-245

125

#domain signal sequence #status experimental #label SIG\

#product insulin-like growth factor-binding protein 4

#status experimental #label MAT\

#domain thyroglobulin type I repeat homology #label THY1\

#binding site carbohydrate (Asn) (covalent) #status predicted

SUMMARY #length 254 #molecular-weight 27745 #checksum 4589

Query Match 4.9%; Score 133; DB 14; Length 254;

Best Local Similarity 39.7%; Pred. No. 1.32e-06;

Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6;

Db 1 mlpfglvaalllaagprpslqde-ahcpcpcseeklarcprpvgeelvrpgcgccatc 59
| | : ||| :|||:|:|:| | :|||:|:|:| | :|||:|:|:| |

Qy 6 MGPRVAFVVLALCSFPAVGQCSGFCR-CPDEPAPRC-PA-GVS-LVLD-GCGCCRCV 60

Db 60 alglgmpe 67
| | | |

Qy 61 AKQLGELC 68

RESULT 9

ENTRY I48599 #type complete

TITLE insulin-like growth factor binding protein 4 - mouse

ORGANISM #formal name Mus musculus #common name house mouse

DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Oct-1996

ACCESSIONS I48599; S38668

REFERENCE I48599

#authors Bethel, C.R.; Vitullo, J.C.; Miller, R.E.; Aron, D.C.

#journal Biochem. Mol. Biol. Int. (1994) 34:385-392

#title Molecular cloning of mouse insulin-like growth factor binding protein 4 (IGFBP4) cDNA and expression of a fusion protein with IGF-binding activity.

#cross-references MUID:95152444

#accession I48599

#status preliminary; translated from GB/EMBL/DBJ

##molecule_type DNA
##residues 1-254 ##label RES
##cross-references EMBL:X76066; NID:g416032; CDS_PID:g416033
CLASSIFICATION #superfamily thyroglobulin type I repeat homology
FEATURE 170-245
SUMMARY #domain thyroglobulin type I repeat homology #label THY1
#length 254 #molecular-weight 27807 #checksum 3509
Query Match 4.9%; Score 133; DB 14; Length 254;
Best Local Similarity 39.7%; Pred. No. 1.32e-06;
Matches 27; Conservative 13; Mismatches 22; Indels 6; Gaps 6;
Db 1 mlpfglvaallaaagrpalsgde-ahcpccseeklarcpvpgceelvrepgcccatc 59
Qy 6 MGPVRVAVVLLALCSRPVQNGSGPCR-CPDEPAPRC-PA-GVS-IVLD-GCGCCRCVC 60
Db 60 alglgmpc 67
Qy 61 AKQLGELC 68

RESULT 10
ENTRY Jc4584 #type complete
TITLE insulin-like growth factor binding protein-5 - pig
ORGANISM #formal name Sus scrofa domestica #common name domestic pig
DATE 10-Apr-1996 #sequence_revision 24-May-1996 #text_change 03-Jun-1996
ACCESSION Jc4584
REFERENCE Jc4584
#authors White, M.E.; Diao, R.; Hathaway, M.R.; Mickelson, J.; Dayton, W.R.
#journal Biochem. Biophys. Res. Commun. (1996) 218:248-253
#title Molecular cloning and sequence analysis of the porcine insulin-like growth factor binding protein-5 complementary deoxyribonucleic acid.

#accession Jc4584
##molecule_type mRNA
##residues 1-271 ##label WHI
##cross-references GB:U41340
#experimental_source skeletal muscle
COMMENT This protein has essential roles in the regulation and coordination of insulin-like growth factors action. This protein enhances the in vitro activity of the insulin-like growth factors, plays a role during myoblast proliferation and differentiation, and is important in the growth and development of muscle tissue. differentiation; growth factor; skeletal muscle

KEYWORDS
FEATURE 1-19
20-271
SUMMARY #domain signal sequence #status predicted #label SIG
#product insulin-like growth factor binding protein-5
#status predicted #label MAT
#length 271 #molecular-weight 30323 #checksum 809

Query Match 4.7%; Score 128; DB 13; Length 271;
Best Local Similarity 37.5%; Pred. No. 8.02e-06;
Matches 30; Conservative 11; Mismatches 31; Indels 8; Gaps 7;
Db 2 vltavilllaacagpagglsfwhcepcdekalmcpcpsplgcelykdpqcgccmtcala 61
Qy 9 VRVAVVLLALCSRPVQNGSGPCR-CPDEPAPRC-PA-GVS-IVLD-GCGCCRCVAKQ 63

Db 62 egsgcgvtyerc-aq-glrc 79
Qy 64 LGELCTE-RDPCDPHKGFLFC 82

RESULT 11
ENTRY Jc1463 #type complete
TITLE insulin-like growth factor-binding protein 5 precursor - rat
ORGANISM #formal name Rattus norvegicus #common name Norway rat
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-1995
ACCESSION Jc1463; A40403; F40403
REFERENCE Jc1463
#authors Zhu, X.; Ling, N.; Shimasaki, S.
#journal Biochem. Biophys. Res. Commun. (1993) 190:1045-1052
#title Cloning of the rat insulin-like growth factor binding protein-5 gene and DNA sequence analysis of its promoter region.

#accession Jc1463
##molecule_type DNA
##residues 1-271 ##label ZHU
##cross-references GB:L08275
REFERENCE A40403
#authors Shimasaki, S.; Shimonaka, M.; Zhang, H.P.; Ling, N.
#journal J. Biol. Chem. (1991) 266:10646-10653
#title Identification of five different insulin-like growth factor binding proteins (IGFBPs) from adult rat serum and molecular cloning of a novel IGFBP-5 in rat and human.

#cross-references M01D:91244847
#accession A40403
##molecule_type mRNA
##residues 1-271 ##label SHI
##cross-references GB:M62781
#accession F40403
#status preliminary
#molecule_type protein
#residues 20-25,'X',27-28,'X',30-36,'X',38-43,'X',45-51,'X',53,'XX' ##label SH2

GENETICS 112/3; 188/3; 228/3
#introns
CLASSIFICATION #superfamily thyroglobulin type I repeat homology
FEATURE 1-19
20-271

#domain signal sequence #status predicted #label SIG
#product insulin-like growth factor binding protein 5
#status predicted #label MAT
#domain thyroglobulin type I repeat homology #label
#length 271 #molecular-weight 30298 #checksum 1075

191-262
SUMMARY
Query Match 4.6%; Score 125; DB 14; Length 271;
Best Local Similarity 36.0%; Pred. No. 2.33e-05;
Matches 27; Conservative 12; Mismatches 28; Indels 8; Gaps 7;

Db 7 llllaacavpagglsfwhcepcdekalmcpcpsplgcelykdpqcgccmtcalaegqsc 66
Qy 14 VVLLALCSRPVQNGSGPCR-CPDEPAPRC-PA-GVS-IVLD-GCGCCRCVAKQLGELC 68
Db 67 gytyerc-aq-glrc 79
Qy 69 TE-RDPCDPHKGFLFC 82

```

RESULT 12
ENTRY   JP0076      #type complete
TITLE   nel protein - chicken
ORGANISM #formal name Gallus gallus #common name chicken
DATE    24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change
                                01-Dec-1995
ACCESSIONS A38963; JP0076
REFERENCE  A38963
#authors   Matsubashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.;
#journal   Taniguchi, S.; Hori, K.
#title     Dev. Dyn. (1995) 203:212-222
#accession A38963
#molecule_type mRNA
#residues  1-835 #label MAT
#cross-references DBJ:D43365
#experimental_source 9-day embryo
REFERENCE  JP0076
#authors   Matsubashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.;
#submision submitted to JPIID, January 1995
#description A new gene, nel, encoding a Mr 93K protein with EGF-like
repeats is strongly expressed in neural tissues of early
stage chick embryos.
#accession JP0076
#molecule_type mRNA
#residues  1-835 #label MA2
#cross-references DBJ:D43365
#experimental_source 9-day embryo
CLASSIFICATION #superfamily von Willebrand factor type C repeat homology
FEATURE
273-333      #domain von Willebrand factor type C repeat homology
#label VWC\
395-592      #region EGF-like repeats
SUMMARY      #length 835 #molecular-weight 93411 #checksum 7565
Query Match 4.6%; Score 126; DB 13; Length 835;
Best Local Similarity 40.3%; Pred. No. 1.63e-05;
Matches 25; Conservative 10; Mismatches 20; Indels 7; Gaps 6;
Db 274 ctmkgmtrefesawtdgck-ncctmngtvqcealcslsdc-p-pnealsy-vc--gkcc 327
Qy 103 CIFGTVYRSGESFSSCKYQCTCIDGAVGCMPL-CSMDVRLPSPDCFFPRVKLPKCC 161
Db 328 ke 329
Qy 162 EE 163

RESULT 13
ENTRY   B37252
TITLE   insulin-like growth factor-binding protein 4 precursor -
human
ALTERNATE_NAMES IGFBP-4; inhibitory insulin-like growth factor-binding
protein; insulin-like growth factor-binding protein pc3
ORGANISM #formal name Homo sapiens #common name man
DATE    28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change

```

```

06-Sep-1996
B37252; B39842; A36549; A60712; A54650; A49801; A34419
REFERENCE  A37252
#authors   Shimasaki, S.; Uchiyama, F.; Shimonaka, M.; Ling, N.
#journal   Mol. Endocrinol. (1990) 4:1451-1458
#title     Molecular cloning of the cDNAs encoding a novel insulin-like
growth factor-binding protein from rat and human.
#cross-references MUID:91133415
#accession B37252
#status    not compared with conceptual translation
#molecule_type mRNA
#residues  1-258 #label SHI
REFERENCE  A39842
#authors   Kiefer, M.C.; Masiarz, F.R.; Bauer, D.M.; Zapf, J.
#journal   J. Biol. Chem. (1991) 266:9043-9049
#title     Identification and molecular cloning of two new 30-kDa
insulin-like growth factor binding proteins isolated from
adult human serum.
#cross-references MUID:91225006
#accession B39842
#molecule_type mRNA
#residues  1-258 #label KIE
REFERENCE  A36549
#authors   Latour, D.; Mohan, S.; Linkhart, T.A.; Baylink, D.J.; Strong,
D.D.
#journal   Mol. Endocrinol. (1990) 4:1806-1814
#title     Inhibitory insulin-like growth factor-binding protein:
cloning, complete sequence, and physiological regulation.
#cross-references MUID:91186988
#accession A36549
#molecule_type mRNA
#residues  1-50, 'A', 52-197, 'F', 199-258 #label LAT
REFERENCE  A60712
#authors   Perkel, V.S.; Mohan, S.; Baylink, D.J.; Linkhart, T.A.
#journal   J. Clin. Endocrinol. Metab. (1990) 71:533-535
#title     An inhibitory insulin-like growth factor binding protein
(in-IGFBP) from human prostatic cell conditioned medium
reveals N-terminal sequence identity with bone derived
in-IGFBP.
#accession A60712
#molecule_type protein
#residues  22-26, 'X', 28-29, 'X', 31-35 #label PER
REFERENCE  A54650
#authors   Mohan, S.; Baylink, D.J.
#journal   Growth Regul. (1991) 1:110-118
#title     Evidence that the inhibition of TE85 human bone cell
proliferation by agents which stimulate cAMP production
in part be mediated by changes in the IGF-II regulatory
system.
#cross-references MUID:93091814
#accession A54650
#molecule_type protein
#residues  22-29, 'X', 31-42 #label MOH
#cross-references NCBI:121076
#experimental_source TE85 osteosarcoma cells
#note      sequence extracted from NCBI backbone
REFERENCE  A49801
#authors   Culouscou, J.M.; Shoyab, M.
#journal   Cancer Res. (1991) 51:2813-2819
#title     Purification of a colon cancer cell growth inhibitor and its

```


	Db	1	mipcalvaallltagprpslgsde-ahncppcseeklarcpvpvgceelvrpgcgcaatc	59
			: : : : : : : :	
	QY	6	MGFVRATVWLLLCSPFAVGNCSGFCR-CPDEPAPRC-PA-GVS-IIVLD-GGCGCRVC	60
			: : : : : : : :	
	Db	60	aiglg	64
	QY	61	AKQLG	65

Search completed: Wed Sep 17 09:30:07 1997
Job time : 76 secs.

maryh@stic

4384-1

NeWSprinter20

Wed Sep 17 11:22:18 1997

NeWSprint 2.5 Rev B

Openwin library 3

NeWSprint interpreter 210.0

NeWSprint 2.5

MPERCH_PP

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelligentGenetics, Inc.

MPerch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 17 09:27:42 1997; MacPar time 9.18 Seconds
Tabular output not generated. 806.556 Million cell updates/sec

Title: >US-08-386-680-2
Description: (1-349) from 5985270.ppep
Perfect Score: 2713
Sequence: 1 MTAASMGPRVAFVLLALC.....PGNDIFESLYRKMGDMA 349

Scoring table: PAM 150
Gap 11

Searched: 59021 seqs, 21210368 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot34
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 44.133; Variance 66.349; scale 0.665

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	2713	100.0	349	3	CTGF HUMAN	0.00e+00
2	2514	92.7	348	3	CTGF MOUSE	0.00e+00
3	1478	54.5	343	6	NOV_XENLA	0.00e+00
4	1460	53.8	351	6	NOV_CHICK	0.00e+00
5	1450	53.4	353	6	NOV_COTJA	0.00e+00
6	1415	52.2	357	6	NOV_HUMAN	0.00e+00
7	1152	42.5	375	2	CE10 CHICK	6.10e-277
8	1113	41.0	379	3	CYR61 MOUSE	4.24e-266
9	133	4.9	254	5	IBP4_RAT	5.18e-09

10	128	4.7	867	9	SSPO_BOVIN	SCO-SPONDIN (FRAGMENT	4.95e-08
11	125	4.6	271	5	IBP5_RAT	INSULIN-LIKE GROWTH F	1.88e-07
12	123	4.5	258	5	IBP4_HUMAN	INSULIN-LIKE GROWTH F	4.54e-07
13	121	4.5	455	1	60A_DROME	60A PROTEIN PRECURSOR	1.09e-06
14	116	4.3	254	5	IBP4_MOUSE	INSULIN-LIKE GROWTH F	9.35e-06
15	113	4.2	272	5	IBP1_RAT	INSULIN-LIKE GROWTH F	3.32e-05
16	113	4.2	2813	10	VWF_HUMAN	VON WILLEBRAND FACTOR	3.32e-05
17	110	4.1	263	5	IBP1_BOVIN	INSULIN-LIKE GROWTH F	1.16e-04
18	111	4.1	271	5	IBP5_MOUSE	INSULIN-LIKE GROWTH F	7.65e-05
19	111	4.1	304	5	IBP2_RAT	INSULIN-LIKE GROWTH F	7.65e-05
20	111	4.1	305	5	IBP2_MOUSE	INSULIN-LIKE GROWTH F	7.65e-05
21	110	4.1	2274	6	MUC2_HUMAN	MUCIN 2 (INTESTINAL M	1.16e-04
22	109	4.0	272	5	IBP1_MOUSE	INSULIN-LIKE GROWTH F	1.75e-04
23	106	3.9	258	5	IBP4_BOVIN	INSULIN-LIKE GROWTH F	5.92e-04
24	105	3.9	272	5	IBP5_HUMAN	INSULIN-LIKE GROWTH F	8.86e-04
25	106	3.9	807	4	FSPO_RAT	F-SPONDIN PRECURSOR.	5.92e-04
26	102	3.8	259	5	IBP1_HUMAN	INSULIN-LIKE GROWTH F	2.92e-03
27	102	3.8	291	5	IBP3_BOVIN	INSULIN-LIKE GROWTH F	2.92e-03
28	104	3.8	311	5	IBP2_CHICK	INSULIN-LIKE GROWTH F	1.32e-03
29	103	3.8	328	5	IBP2_HUMAN	INSULIN-LIKE GROWTH F	1.97e-03
30	103	3.8	500	10	VLI_HPV6B	MAJOR CAPSID PROTEIN	1.97e-03
31	102	3.8	798	5	ITB7_HUMAN	INTEGRIN BETA-7 SUBUN	2.92e-03
32	104	3.8	3133	5	HMCYT_BOVMO	HEMOCYTIN PRECURSOR (1.32e-03
33	101	3.7	388	3	CSP_FLARE	CIRCUMSPOROZOITE PROT	4.32e-03
34	101	3.7	399	1	EM8A_MOUSE	BONE MORPHOGENETIC PR	4.32e-03
35	101	3.7	402	1	EMP8_HUMAN	BONE MORPHOGENETIC PR	4.32e-03
36	101	3.7	405	2	CSP_PLAFO	CIRCUMSPOROZOITE PROT	4.32e-03
37	101	3.7	412	2	CSP_PLAFA	CIRCUMSPOROZOITE PROT	4.32e-03
38	101	3.7	424	3	CSP_PLAFT	CIRCUMSPOROZOITE PROT	4.32e-03
39	100	3.7	442	3	CSP_PLAFW	CIRCUMSPOROZOITE PROT	6.38e-03
40	101	3.7	551	11	YVD3_CAEEL	HYPOTHETICAL ZINC MET	4.32e-03
41	99	3.6	111	5	IBP5_BOVIN	INSULIN-LIKE GROWTH F	9.40e-03
42	99	3.6	291	5	IBP3_MOUSE	INSULIN-LIKE GROWTH F	9.40e-03
43	99	3.6	292	5	IBP3_RAT	INSULIN-LIKE GROWTH F	9.40e-03
44	98	3.6	806	5	ITB7_MOUSE	INTEGRIN BETA-7 SUBUN	1.38e-02
45	96	3.5	803	4	FSPO_XENLA	F-SPONDIN PRECURSOR.	2.95e-02

ALIGNMENTS

RESULT	1	ID	CTGF HUMAN	STANDARD;	PRT;	349 AA.
AC	P29279;					
DT	01-DEC-1992		(REL. 24, CREATED)			
DT	01-DEC-1992		(REL. 24, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996		(REL. 34, LAST ANNOTATION UPDATE)			
DE	CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.					
CN	CTGF.					
OS	HOMO SAPIENS (HUMAN).					
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;					
OC	EUTHERIA; PRIMATES.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;					
PX	MEDLINE; 91373462.					
RA	BRADHAM D.M., IGARASHI A., POTTER R.L., GROTEENDORST G.R.;					
RL	J. CELL BIOL. 114:1285-1294 (1991).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;					

Wed Sep 17 11:16:43 1997

Listing for Mary Hale

RX MEDLINE; 93187114.
 RA IGARASHI A., BRADHAM D.M., OKOCHI H., GROTEENDORST G.R.;
 RL J. DERMATOL. 19:642-643(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA OSMAR B.S., WERNER A., YANG Z., GARNIER J.M., GENTZ R., LUESCHER T.F.;
 RL SUBMITTED (APR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
 CC HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN
 CC MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.
 CC -1- SUBUNIT: MONOMER.
 CC -1- ALTERNATIVE PRODUCTS: A SHORTER FORM MAY BE PRODUCED BY
 CC ALTERNATIVE SPLICING OF THE SAME GENE.
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; M292934; G180924; -;
 DR EMBL; M56201; G266235; -;
 DR EMBL; X78947; G474934; -;
 DR PIR; A40531; A40531.
 DR PIR; S44203; S44203.
 DR MIM; 121009; -;
 DR PROSITE; PS00222; IGF BINDING.
 KW GROWTH FACTOR BINDING; SIGNAL; ALTERNATIVE SPLICING.
 FT SIGNAL 1 21
 FT CHAIN 22 349
 FT DOMAIN 101 167
 FT DOMAIN 256 330
 FT DISULFID 256 293
 FT DISULFID 273 307
 FT DISULFID 284 323
 FT DISULFID 287 325
 FT DISULFID 292 329
 FT CARBOHYD 28 28
 FT CARBOHYD 225 225
 FT VARSPLIC 172 198
 FT SEQUENCE 349 AA; 38069 MW; C2IE9662 CRC32;
 Query Match 100.0%; Score 2713; DB 3; Length 349;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 mtaasgprvrvafvllalcsrpavqncsgpcrcpdpaprcpavslvdgcgcrcv 60
 QY 1 MTAASGPRVRFVLLALCSRPVAVQNCSGPCRCDEPAPRCAPAGVSLVLDGCGCRVC 60
 Db 61 akqlgelcterdpchkgllfcdgspanrkigvctakdgapcifggtvyrsgesfgsc 120
 QY 61 AKQGLGELCTERDPCPHKGLFCDGSPANRKIGVCTAKDGAPCIFGGTVYRSGESFGSSC 120
 Db 121 kvqctldgavcgmplcsmdvrlpsdpcpfprvklpgkccceewvcddepkqdtvvgpala 180
 QY 121 KVQCTCLDGA VGMPLCSMDVRLPSDPCPFPRVKLPKGCCEEWVCDDEPKQDQTVVGPALA 180
 Db 181 avrledtfgdpdmiranclvqtetwsacsktcgmigistrvndnasrleqkqrlcmvr 240
 QY 181 AVRLEDTFGDPDMIRANCLVQTEWSACSKTCGMIGISTRVNDNASRLEKQRLCMVR 240
 Db 241 pceadleenikkkgkciirtpkiskpikfelagctamtktyrakfgvctdgrcctphrttt 300

Wed Sep 17 11:16:43 1997

Listing for Mary Hale

QY 241 PCEADLEENIKKKKCIIRTPKISKPIKFELSGCTSMKTYRAKFCGCTDRCCTPHRTTT 300
 Db 301 lpvefkcpdgvemkmmfiktacachyncpgndndifesyryrmygdma 349
 QY 301 LPVEFKCPDGEVMKMMFIKTCACHYNCPGNDNDFESLYRYRMYGDMA 349
 RESULT 2
 ID CTGF MOUSE STANDARD; PRT; 348 AA.
 AC P29268;
 DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (CTGF) (FISP-12 PROTEIN).
 GN CTGF OR FISP12 OR FISP-12.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91363290.
 RA RYSECK R.-P., MACDONALD-BRAVO H., MATTEI M.-G., BRAVO R.;
 RL CELL GROWTH DIFFER. 2:225-233(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91229699.
 RA BRUNNER A., CHINN J., NEUBAUER M.G., PURCHIO A.F.;
 RL DNA CELL BIOL. 10:293-300(1991).
 CC -1- INDUCTION: BY GROWTH FACTORS.
 CC -1- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN
 CC (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A VWFC DOMAIN.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; M70641; G193314; -;
 DR EMBL; M70642; G193316; -;
 DR EMBL; M80263; G201946; -;
 DR PIR; A53228; A53228.
 DR PROSITE; PS00222; IGF BINDING.
 KW GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 21
 FT CHAIN 22 348
 FT DOMAIN 100 166
 FT DOMAIN 255 329
 FT DISULFID 255 292
 FT DISULFID 272 306
 FT DISULFID 283 322
 FT DISULFID 286 324
 FT DISULFID 291 328
 FT CONFLICT 161 161
 FT SEQUENCE 348 AA; 37793 MW; EAB92BEO CRC32;
 Query Match 92.7%; Score 2514; DB 3; Length 348;
 Best Local Similarity 90.8%; Pred. No. 0.00e+00;
 Matches 317; Conservative 21; Mismatches 10; Indels 1; Gaps 1;
 Db 1 mlaavagpisialv-llalctrpatgdcacqcaaaaphcpagvslvdgcgcrcv 59

QY 1 MTAASMGPRVAVFVLLALCSRPAGVQNCSPCRCPDEPAPCPAGVSLVLDGCGCCRV 60
 Db 60 akqlgelcterdpdphkglfcdgfgspanrkigvctakdgapcvfggavvragesfqsc 119
 QY 61 AKQLGELCTERDPDPCDHPKGLFCDFGSPANRKIGVCTAKDGAPCIGFTVYRSGESFQSSC 120
 Db 120 kygcetldgavgcplcsmdvrlpsdpcpfrvrk1pgkckewcdkdpkrtavgpala 179
 QY 121 KYQCTCLDGAAGVGMFLCSMDVRLPSDPCPFRVRK1PGKCCSEWVCEDEPKDTVVGPALA 180
 Db 180 ayrltdtfgpdtmttrancvlqtewasactgcmgistrvndntforlekqsrclmvr 239
 QY 181 AYRLDTFGPDTMIRANCLVQTEWSACSKTCGMGISTRVNDNASCRLEKQSRCLMVR 240
 Db 240 pceadleenikgkckirtpkiaqvkfclsgctsvktyrakfcgvcctdgrcctphrttt 299
 QY 241 PCEADLEENIKGKCKIRTPKISKPIKFELSGCTSMKTYRAKFCGVCCTDGRCTPHRTTT 300
 Db 300 lpvefkcpgdgeincknmfiktcaachynopgdndifaellyrkmvgdina 348
 QY 301 LPVEFKCPDGEVMMKNMFIKTCACHYNCFGDNDIFESLYRKMVGDMA 349

RESULT 3 STANDARD; PRT; 343 AA.

AC P51609;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DE NOV PROTEIN HOMOLOG PRECURSOR (XNOV).
 GN NOV
 OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA YING Z., KING M.L.;
 RL SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 GROWTH REGULATION (BY SIMILARITY).
 CC -!- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -!- SIMILARITY: CONTAINS A VWFC DOMAIN.
 CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; U37063; G1176415; -.
 KW GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 18
 FT CHAIN 19 343
 FT DOMAIN 93 159
 FT DOMAIN 249 323
 FT DISULFID 249 286
 FT DISULFID 266 300
 FT DISULFID 277 316
 FT DISULFID 280 318
 FT DISULFID 285 322
 FT CARBOHYD 265 265
 SQ SEQUENCE 343 AA; 38070 MW; BDF3BCA4 CRC32;

Query Match 54.5%; Score 1478; DB 6; Length 343;
 Beat Local Similarity 53.7%; Pred. No. 0.00e+00;
 Matches 180; Conservative 69; Mismatches 78; Indels 8; Gaps 8;

Db 1 mtp-hlalcfill-l-lqqvasqkpcqcdqcpcep-pcacpsvillldgcqccpvcqrq 56
 QY 6 MGPVRVAVFVLLALCSRPAGVQNCSPCRCPDEPAPCPAGVSLVLDGCGCCRVCAKQL 64
 Db 57 gescchlnpcqgedklycefnadprmetgtcmalegnacvfdgvvynreafpsckyh 116
 QY 65 GELCTERDPDPCDHPKGLFCDFGSPANRKIGVCTAKDGAPCIGFTVYRSGESFQSSCKYQC 124
 Db 117 tclnhighcvprcnldllllpgdpdpfrvrkvpgecekwcd-skeemaigfgfanaayr 175
 QY 125 TCLDGAAGVGMFLCSMDVRLPSDPCPFRVRK1PGKCCSEWVCEDEPKDTVVGP-ALAAAYR 183
 Db 176 pcatlidasdstafacttewasactgcmgistrvndntforlekqsrclmvr 235
 QY 184 LEDTFGPDTMIRANCLVQTEWSACSKTCGMGISTRVNDNASCRLEKQSRCLMVRPCE 243
 Db 236 eepgwhvekgkckvrvrkttkpihfyknctsvqpykpkfcgqcdgrcctphaktmh 299
 QY 244 ADLEENI-KGKCKIRTPKISKPIKFELSGCTSMKTYRAKFCGVCCTDGRCTPHRTTLP 302
 Db 296 vefvcpqkrivkvpvmvstcvcvhynpcqdsllq 330
 QY 303 VEFKCPDGEVMMKNMFIKTCACHYNCFGDNDIFE 337

RESULT 4

ID NOV CHICK STANDARD; PRT; 351 AA.
 AC P28686;
 DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DE NOV PROTEIN PRECURSOR.
 GN NOV
 OS GALLUS GALLUS (CHICKEN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
 RN GALLIFORMES.
 RP SEQUENCE FROM N.A.
 RC STRAIN=BROWN LEGHORN;
 RX MEDLINE; 92107157.
 RA MALOISEL V., MARTINIE C., DAMBRINE G., PLASSIART G., BRISAC M.,
 RA CROCHET J., PERBAL B.;
 RL MOL. CELL. BIOL. 12:10-21(1992).
 CC -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH
 TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION
 OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT
 TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.
 CC -!- DEVELOPMENTAL STAGE: MAVI-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH
 LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN
 ADULT KIDNEY.
 CC -!- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL IN
 MUSCLE AND INTESTINE, IN THE EMBRYO. LUNG AND LESS SO IN BRAIN AND
 SPLEEN, IN ADULT CHICKEN.
 CC -!- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -!- SIMILARITY: CONTAINS A VWFC DOMAIN.
 CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; X59284; G63703; -.

DR	PIR; S20078; S20078.	DR	PROSITE; PS00222; IGF BINDING.
KW	PROTO-ONCOGENE; GROWTH FACTOR BINDING; SIGNAL.	KW	PROTO-ONCOGENE; GROWTH FACTOR BINDING; SIGNAL.
FFT	SIGNAL	FFT	SIGNAL
1	24	1	24
FFT <td>CHAIN</td> <td>FFT<td>CHAIN</td></td>	CHAIN	FFT <td>CHAIN</td>	CHAIN
25	351	25	351
DOMAIN	104	DOMAIN	104
258	332	258	332
FFT <td>DOMAIN</td> <td>FFT<td>DOMAIN</td></td>	DOMAIN	FFT <td>DOMAIN</td>	DOMAIN
258	295	258	295
DISULFID	275	DISULFID	275
309	309	309	309
FFT <td>DISULFID</td> <td>FFT<td>DISULFID</td></td>	DISULFID	FFT <td>DISULFID</td>	DISULFID
286	325	286	325
FFT <td>DISULFID</td> <td>FFT<td>DISULFID</td></td>	DISULFID	FFT <td>DISULFID</td>	DISULFID
289	327	289	327
FFT <td>DISULFID</td> <td>FFT<td>DISULFID</td></td>	DISULFID	FFT <td>DISULFID</td>	DISULFID
294	331	294	331
FFT <td>CARBOHYD</td> <td>FFT<td>CARBOHYD</td></td>	CARBOHYD	FFT <td>CARBOHYD</td>	CARBOHYD
274	274	274	274
SEQUENCE	351 AA;	SEQUENCE	351 AA;
38268 MW;	C70404065 CRC32;	38268 MW;	C70404065 CRC32;

Query Match 53.8%; Score 1460; DB 6; Length 351;
Best Local Similarity 54.2%;
Pred. No. 0.00e+00;
Matches 186; Conservative 59; Mismatches 80; Indels 8; Gaps 7;

Db		3	tgggqglpvlllllllrrpoevsgreaacprpcggrcraep-prcapgpvavldgcgccl	61
Qy		2	TAA\$MG--PVRVAFVWLLALCSRPAGVQNGSGPC--RCDEPAAPRCPAGSVLVLDDGGCCCR	58
bB		62	vcarqrgeescpllpodeagglvcdrgrpedgggagicmvlegdncvfddgmlyrngctfcp	121
Qy		59	VCAQLGELCTERDPCDPHHKGLFCDFGSPANRKIGVCTAKDGAPCFICFVGTVYRSGESFO	118
bB		122	ackvqtcdrqdgigelpcnlgllpgdpdcfpkrieypcecekwwcd-prdevillg	180
Qy		119	SKYQCTCLDGAVCGMPLSCMDVRFLPSDFCFFPRVKLPFGKCCEEWCDPEKDQTVVGP-	177
bB		181	amaayreatlgidvadseanciedtewsacskschgmgfstvtvtnnqqcmvkqtrlc	240
Qy		178	ALAAYLEDTFGDPTMIRANCLVQTTEWSACSKTCGMGISTRVNDNASCLEXQSRLC	237
bB		241	mnrpcone-epedckgkciqtckkamkarvfeykntcsavtykpycgldndgroctphn	299
Qy		238	MVRPCEADLEENIKKKCI RTPKTSKP IKFSLGCTSMKTYRAKTGCVGTDCRGCTPHR	297
bB		300	tktiqvfcrcpqgklfkkmmlintcvchncpnagnaffql	342
Qy		298	TTLTPVEFKCPDGEVMKKNNMF IKTCACHNCP-GDNDIFESL	339

RESULT	5	STANDARD;	PRT;	353 AA.
D NOV COTJA				
C P22302				
WT 01-NOV-1995		(REL. 32, CREATED)		
WT 01-NOV-1995		(REL. 32, LAST SEQUENCE UPDATE)		
WT 01-OCT-1996		(REL. 34, LAST ANNOTATION UPDATE)		
WT 01-OCT-1996		NOV PROTEIN PRECURSOR.		
WT 01-OCT-1996				

COTURNIX COTURNIX JAPONICA (JAPANESE QUAIL).
EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
CALLIFORMES.
[1]
SEQUENCE FROM N.A.
WEISKIRCHEN R., BISTER K.;
SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
-1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL

-1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL

CC	GROWTH REGULATION (BY SIMILARITY).
CC	-!- SIMILARITY TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
CC	TO THE CEF-10/CYR61/CTFG/FIS-12/NOV PROTEIN FAMILY.
CC	-!- SIMILARITY: CONTAINS A VWFC DOMAIN.
CC	-!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC	EMBI: U13063; G532697; ...
DR	PROSITE; PS00222; IGF_BINDING.
DR	PROTO-ONCOGENE; GROWTH FACTOR BINDING; SIGNAL.
KW	POTENTIAL.
FT	CHAIN 1 26
FT	CHAIN 27 353
FT	DOMAIN 106 172
FT	DOMAIN 260 334
FT	DOMAIN 260 334
FT	DISULFID 260 297
FT	DISULFID 277 311
FT	DISULFID 288 327
FT	DISULFID 291 329
FT	DISULFID 296 333
FT	CARBOHYD 276 276
FT	POTENTIAL.
FT	SEQUENCE 353 AA; 386667 MW; C4F5928D CRC32;
CSQ	

Query Match 53.4%; Score 1450; DB 6; Length 353;
Best Local Similarity 54.7%;
Pred. No. 0.00e+00;
Matches 185; Conservative 66; Mismatches 78; Indels 9; Gaps 8;

10 pvvlllllllllllpsevnreaprcprrcggrcpaep-prcapgypavldgcgcclvcraq 68
8 PVRVAFVULLALCSREAVG-QN-C5GPC--RCDEPAPRCPAGVSLVLDGCGGCCRCVCAQ 63
69 rgescpllpdcsegglyodrpdedgggtgicmwlegdcvfdgmilyngctfpapcq 128
64 LGELCTEROPCDPHKGLGDFGSPANRKIGVCTAKDAPCIFGTVTVRSGESFQSSCKY 123
129 ctcrdggicqlprcnllgpdcpcpfrkievgececekwc-eprdevllvgfamaay 187
124 CTCLDGVAGGMPJCSMDVRLSPDFPFRVKLPJGKCEEWVDEPKDQTVVP-ALAA 182
188 rgeatlgidvadesancieqtseasarscgmfgstrvnrnqgcmvkgqtrclmmrcp 247
183 RLEDYTGPDPTMRANCLVQTTWESACSKTGMGISTRVNDNASCRLEKQSLCMLVRPC 242
248 ene-epesdkkgkcirtkksmkaufeykntesvqtkpvcyglcndgcrctphntktiq 306
243 EADLEENIKKGKKGKIRTPKISKPIKELSGCTSMKTYRAKFGCVTDGRCCTPHRTTLP 302
307 vefrcpgkflkkmmlintvcvchncpqsnnaffqpl 344
303 VEFKCPDGEVMKKNMWFKTCACHYVCP-GNDIFESL 339

RESULT		6		STANDARD;		PRT;		357 AA.	
ID	NOV HUMAN								
AC	248745;								
TT	01-FEB-1996			(REL. 33, CREATED)					
TT	01-FEB-1996			(REL. 33, LAST SEQUENCE UPDATE)					
TT	01-OCT-1996			(REL. 34, LAST ANNOTATION UPDATE)					
DE	NOV PROTEIN			HOMOLOG PRECURSOR (NOVH).					

OS
OC
OC

RV SEQUENCE FROM N.A.
 RP TISSUE=PLACENTA;
 RX MEDLINE; 94336229.
 RA MARTINERIE C., HOFF V., JOUBERT I., BADZIOCH M., SAUNDERS G.,
 RL ONCOGENE 9:2729-2732(1994).
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC -1- GROWTH REGULATION (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: INCREASED EXPRESSION IN WILMS TUMOR OF THE
 CC STROMAL TYPE.
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A VFMC DOMAIN.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; X78351; G825696; --.
 DR EMBL; X78352; G825696; JOINED.
 DR EMBL; X78353; G825696; JOINED.
 DR EMBL; X78354; G825696; JOINED.
 DR EMBL; X96584; E228691; --.
 DR MTM; 164958; --.
 KW PHOTO-ONCOGENE; GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 27
 FT CHAIN 28 357
 FT DOMAIN 108 174
 FT DOMAIN 264 338
 FT DISULFID 264 301
 FT DISULFID 281 315
 FT DISULFID 292 331
 FT DISULFID 295 333
 FT DISULFID 300 337
 FT CARBOHYD 97 97
 FT CARBOHYD 280 280
 FT SEQUENCE 357 AA; 39162 MW; DAB8009D CRC32;
 Query Match 52.2%; Score 1415; DB 6; Length 357;
 Best Local Similarity 51.8%; Pred. No. 0.00e+00;
 Matches 177; Conservative 68; Mismatches 91; Indels 6; Gaps 5;
 Db 18 ltfll-llhllgvaatqrcppqgpcratp-ptcagpvtavldgcacclivcarqtgsc 75
 Qy 11 VAFVLLALCSRPAGVQNCSPG--RCPDEPAPRCAGVSLVLDGCGCCRVCAKQJGELC 68
 Db 76 adlpcdesaglycdreadpanqdgictavegdnvcvfdgviyrsgkfqsckfqtcd 135
 Qy 69 TERDPCDHPKGLFCDFGSPANRKIGVCTAKDCAFCIFGTVYRSGESFQSSCKYQCTCLD 128
 Db 136 gqigcvprcdlvdllpepncpaprkvvevpgeccckwldgdeadalggltlaayrpeatl 195
 Qy 129 GAVGCMPLCSMDVRLSPDCDFPFRVRLPKGKCEEWCDPEKDTQVVGPAALAYRLEDTF 188
 Db 196 gvevadsavncieqtetwtacskacgmfgfstrvtnnrqcmikqtrlcmvrpcegepeq 255
 Qy 189 GPDEPTMIRANLVQTEWSACSKTCGMGISTRVNDNASCRLEKQSRLCWVRECEADLEE 248
 Db 256 ptdkkgkclrtkshlkhqkncstlhtyprfgovsdgrcctphntkqieafgc 315
 Qy 249 NI-KKGGKCIINTPKISKPIKELSGCTSMKTYRAKFCGVCTDGRCTPHRTTTLPEFKC 307
 Db 316 adqigvkvpmvigtctctchncpknneafqlqlelkttrckm 357

QY 308 PDGEVNMKNMFIKTCAHCYNGPNDIF-ESLYYRKMYGDM 348
 RESULT 7
 ID CE10 CHICK STANDARD; PRT; 375 AA.
 AC P19336;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE CEF-10 PROTEIN PRECURSOR.
 OS GALLUS GALLUS (CHICKEN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
 OC GALLIFORMES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89145206.
 RA SIMMONS D.L., LEVY D.B., YANNONI Y., ERIKSON R.L.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:1178-1182(1989).
 CC -1- FUNCTION: PROBABLE SECRETED REGULATORY PROTEIN.
 CC -1- INDUCTION: BY V-SRC.
 CC -1- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A VFMC DOMAIN.
 CC -1- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; J04496; G211436; --.
 DR PIR; A41428; A41428.
 DR PROSITE; PS00222; IGF BINDING.
 KW GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 22
 FT CHAIN 23 375
 FT DOMAIN 98 164
 FT DOMAIN 281 355
 FT DISULFID 281 318
 FT DISULFID 298 332
 FT DISULFID 309 348
 FT DISULFID 312 350
 FT DISULFID 317 354
 FT SEQUENCE 375 AA; 40651 MW; 68B4BC92 CRC32;
 Query Match 42.5%; Score 1152; DB 2; Length 375;
 Best Local Similarity 48.5%; Pred. No. 6.10e-277;
 Matches 176; Conservative 76; Mismatches 77; Indels 34; Gaps 19;
 Db 7 rpalaal-lclarlalagpcpavacqpa-aapqcapvglyvdpdgcgcckvcakqlnedc 64
 Qy 10 RVAFVLLALC-SRPAGVQNCSPGCRCPDEPAPRCAGVSLVLDGCGCCRVCAKQJGELC 68
 Db 65 artqpcdhtkglcfnfgaspaatngicraqsegrpceynaskiyngesfqpnckhqtcti 124
 Qy 69 TERDPCDHPKGLFCDFGSPANRKIGVCTAK-DGAPCIFGTVYRSGESFQSSCKYQCTCL 127
 Db 125 dgavgcipcpqelalpnlcgpcprlvkvpgccceewvcdskdalelegffakefgld 184
 Qy 128 DGAVGCMPLCSMDVRLSPDCDFPFRVRLPKGKCEEWCDPEKD--QTVVG--P---AL- 179
 Db 185 aasegeltrnnelalvkggklmfpvfgaepqsrarfenpkciqvttswsqscaktgtglat 244
 Qy 180 AA---Y-R---L-E-----D-----TFGDPD-TM-I-RANCLVQTTSACSKTCGMGIST 219

Db 245 rvtndpckliketricevzpcgpayasalkkgtkcttkkpsvrfyagcsavkky 304
 Qy 220 RVTNDNASCRLKOSRLCMVRPCBADLEENIKRGRKCIPTPKIKFELSGCTSMKTY 279
 Db 305 rpykycgscvdcrcctpcgttvtkirkfcdcdggetftkavmmigecrcnynchaneay-p- 362
 Qy 280 RAKFCGCTDGRCTPHTTTLVPEFKCPDGEVKKMMFIKTACHYNCPCDNDIFESL 339
 Db 363 fyr 365
 Qy 340 YR 342

RESULT 8
 ID CYR6 MOUSE STANDARD; PRT; 379 AA.
 AC P18406;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE CYR61 PROTEIN PRECURSOR (3CH61).
 GN CYR61.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C / 3T3;
 RX MEDLINE; 90287146.
 RA O'BRIEN T.P., YANG G.P., SANDERS L., LAU L.F.;
 RL MOL. CELL. BIOL. 10:3569-3577(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AJ; TISSUE=EMBRYONIC FIBROBLAST;
 RX MEDLINE; 91288203.
 RA LATINKIC B.V., O'BRIEN T.P., LAU L.F.;
 RL NUCLEIC ACIDS RES 19:3261-3267(1991).
 CC -|- FUNCTION: CYR61 MIGHT ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS.
 CC -|- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN
 CC NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
 CC -|- INDUCTION: BY GROWTH FACTORS.
 CC -|- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
 CC AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST
 CC IN LUNG.
 CC -|- SIMILARITY: TO INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS AND
 CC TO THE CEP-10/CYR61/CTFG/FISP-12/NOV PROTEIN FAMILY.
 CC -|- SIMILARITY: CONTAINS A VMEC DOMAIN.
 CC -|- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 DR EMBL; M32490; G309206; -.
 DR PIR; A35669; G50633; -.
 DR PROSITE; PS00222; IGF BINDING.
 KW GROWTH FACTOR BINDING, SIGNAL.
 FT SIGNAL 1 24
 FT CHAIN 25 379
 FT CYR61 PROTEIN.
 FT DOMAIN 98 164
 FT DOMAIN 284 358
 FT CTCK.
 FT DISULFID 284 321
 FT BY SIMILARITY.
 FT DISULFID 301 335
 FT BY SIMILARITY.
 FT DISULFID 312 351

FT DISULFID 315 353 BY SIMILARITY.
 FT DISULFID 320 357 BY SIMILARITY.
 SQ SEQUENCE 379 AA; 41709 MW; 116B80C7 CRC32;
 Query Match 41.0%; Score 1113; DB 3; Length 379;
 Best Local Similarity 48.3%; Pred. No. 4,246-266;
 Matches 180; Conservative 79; Mismatches 76; Indels 38; Gaps 20;

Db 1 masstfrtlava-vtllhl-trlals-tcpaachople-apkcapgvglvrdgcgckvc 56
 Qy 1 MTAASMGVVRVAFVVLALCSRPVAVGQNCSPCPDPAPRCRCPAGVSLVLDGCGCRVC 60
 Db 57 akqlnedcstqpcdhtkglecnfgasatalkgicraqsegprceynsriyngesfgpn 116
 Qy 61 AKQLGELCTERDPCDPHKGFLCDFGSPANRKGIVCTAK-DGAPCIFGTVIRGESFQSS 119
 Db 117 ckhqctcidgavgcipclpcqelslpnlgcpnprlvkvsqgceeswcdedsikdsiddgd 176
 Qy 120 KKYQCTCLDGAVGCMPLCSMDVRLPSPDCPPFRVRLKPLGKCEEWCDCE-P-KD-----QT 1
 Db 177 dilglaseveltrnneliiagksalrkripvfgtprvlfnpahgkqciqvgttaweq 236
 Qy 174 -VVG--PA---LA-----AY-R---LE---DTFGDP-TM---IRAN---CLVQTTEWSA 208
 Db 237 cskscgtgistrvtndnpeclrvketricevrcpgqpyasalkkgtkckskkpsvrf 296
 Qy 209 CSKTGCMGISTRVNDNASCRLKOSRLCMVRPCBADLEENIKRGRKCIPTPKIKF 268
 Db 297 tyagssvkvrykycgscvdcrcctplqtrtvkmrfceedgemfknvmmigscckcyn 356
 Qy 269 ELSGCTSMKTYRAKFCGCTDGRCTPHTTTLVPEFKCPDGEVKKMMFIKTACHYN 328
 Db 357 cphneasf-rly 368
 Qy 329 CPGDNDI-FESLY 340

RESULT 9
 ID IBP4 RAT STANDARD; PRT; 254 AA.
 AC P21744;
 DT 01-MAY-1991 (REL. 18, CREATED)
 DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 PRECURSOR (IGFBP-4)
 DE (IBP-4) (IGF-BINDING PROTEIN 4).
 GN IGFBP4 OR IGFBP-4.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91133415.
 RA SHIMASAKI S., UCHIYAMA F., SHIMONAKA M., LING N.;
 RL MOL. ENDOCRINOL. 4:1451-1458(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE; 93176147.
 RA GAO L., LING N., SHIMASAKI S.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 190:1053-1059(1993).

[3]	SEQUENCE OF 22-61.	
RC	TISSUE-SERUM;	
RR	MEDLINE; 90073708.	
RA	SHIMONAKA M., SHIMROEDER R., SHIMASAKI S., LING N.;	
RL	BIOCHEM. BIOPHYS. RES. COMMUN. 165:189-195(1989).	
CC	-I- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs	
CC	AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH	
CC	PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE	
CC	INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.	
CC	-I- BINDS IGF-II MORE THAN IGF-I.	
CC	-I- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.	
DR	EMBL; L08276; E73337; --.	
DR	PIR; B33570; B33570.	
DR	PIR; A37252; A37252.	
DR	PIR; JC1464; JC1464.	
DR	PIR; E40403; E40403.	
DR	PROSITE; PS00222; IGF BINDING.	
DR	PROSITE; PS00484; THYROGLOBULIN 1.	
KKW	GROWTH FACTOR BINDING; SIGNAL; GLYCOPROTEIN.	
FT	SIGNAL	1 21
FT	CHAIN	22 254
FT	CARBOHYD	125 125
FT	POTENTIAL.	
FT	DOMAIN	196 245
FT	SEQUENCE	254 AA; 27745 MW; 3B5316DC CRC32;
SO	INSULIN-LIKE GROWTH FACTOR BINDING	
	PROTEIN 4.	
	POTENTIAL.	
	THYROGLOBULIN TYPE I.	

	Query Match	4.9%;	Score 133;	DB 5;	Length 254;
	Best Local Similarity	39.7%;	Pred. No. 5.18e-09;		
	Matches	27;	Conservative 13;	Mismatches 22;	Indels 6; Gaps 6;
bb	1	mlpfglvaalllaagprpalgde-ahhpcpseeeklarcrppvgceslvrepqgcgcac	59		
yy	6	MGPRVAFVVLILALCRSPAVGQNGSPCR-CFDEFAPRC-PA-GVS-LVLD-GCGCGRVC	60		
bb	60	alglmpc	67		
yy	61	AKQLGELC	68		

[illegible]

CC	-1-	DEVELOPMENTAL STAGE: EMBRYO.		
CC	-1-	SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.		
CC	-1-	SIMILARITY: CONTAINS AT LEAST 2 EGF-LIKE DOMAINS.		
CC	-1-	SIMILARITY: CONTAINS AT LEAST 3 LDL-RECEPTOR CLASS A DOMAINS.		
DR	EMBL: X93922; E213357; -.			
KW	GLYCOPROTEIN; CELL ADHESION; CALCIUM-BINDING; REPEAT; EGF-LIKE DOMAIN.			
FT	NON TER	1		
FT	DOMAIN	103	142	EGF-LIKE.
FT	DOMAIN	143	180	EGF-LIKE.
FT	DOMAIN	506	544	LDL-RECEPTOR CLASS A.
FT	DOMAIN	663	701	LDL-RECEPTOR CLASS A.
FT	DOMAIN	723	761	LDL-RECEPTOR CLASS A.
FT	CARBOHYD	88	88	POTENTIAL.
FT	CARBOHYD	309	309	POTENTIAL.
FT	CARBOHYD	409	409	POTENTIAL.
FT	NON TER	867	867	POTENTIAL.
SQ	SEQUENCE	867 AA;	91817 MW;	B1224081 CRC32;

Query Match	4.7%	Score 128;	DB 9;	Length 867;
Best Local Similarity	28.3%	Pred. No. 4.95e-08;		
Matches	32;	Conservative	21;	Mismatches 48;
			Indels 12;	Gaps 12;
Db	81	cpdgcgpnvtcsgeglvfhacvcpcltcdtdisgqatcpcpdr-pcggp-gcwcgagqvlgaq	138	
QY	35	CPDEPAPRCFAGVSLVLDGGCCRV-CARQLGEL-CT-ERDPPCHKGLFCDFGSPANRK	91	
Db	139	-grcwprpq-cpelvdgsrywpqgrvktcdql-cvcqdgrrprcqaalcaavn	188	
QY	92	IGVCT-ARKGAPCIFGTVVRSGESFQSSCKRYQCTCLDG-AVGQMP-L-GMSD	140	

RESULT	11
ID	IBP5 RAT
AC	STANDARD; PRT; 271 AA.
P24594;	
DDT	01-MAR-1992 (REL. 21, CREATED)
DDT	01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DDT	01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE	INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 5 PRECURSOR (IGFBP--5)
DE	(IBP-5) (IGF-BINDING PROTEIN 5).
GN	IGFBP5 OR IGFBP-5. (RAT).
OS	RATTUS NORVEGICUS. (RAT).
OCC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OCC	EUTHERIA; RODENTIA.
[1]	
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 20-53.
RP	TISSUE=OVARY;
RC	MEDLINE; 91244847.
RR	SHIMASAKI S., SHIMONAKA M., ZHANG H.-P., LING N.;
RR	J. BIOL. CHEM. 266:10646-10653(1991).
[2]	
RP	SEQUENCE FROM N.A.
RP	STRAIN=SPRAGUE-DAWLEY;
RC	MEDLINE; 93176146.
RR	ZHU X., LING N., SHIMASAKI S.;
RR	BIOCHEM. BIOPHYS. RES. COMMUN. 190:1045-1052(1993).
RL	-!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE I
CC	AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
CC	PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
CC	INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
CC	-!- TISSUE SPECIFICITY: MOSTLY IN KIDNEY.

-!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.

CC EMBL; M62781; G204746; -.
 DR EMBL; L08275; E73333; -.
 DR PIR; A40403; A40403.
 DR PIR; JCI1463; JCI1463.
 DR PIR; F40403; F40403.
 DR PROSITE; PS00222; IGF BINDING.
 DR PROSITE; PS00484; THYROGLOBULIN_1.
 KW GROWTH FACTOR BINDING; SIGNAL.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 271 INSULIN-LIKE GROWTH FACTOR BINDING
 FT CHAIN 20 271 PROTEIN 5.
 FT DOMAIN 214 262 THYROGLOBULIN TYPE 1.
 FT SEQUENCE 271 AA; 30298 MW; 0AA79306 CRC32;

Query Match 4.6%; Score 125; DB 5; Length 271;
 Best Local Similarity 36.0%; Pred. No. 1.88e-07;
 Matches 27; Conservative 12; Mismatches 28; Indels 8; Gaps 7;

Db 7 llllaacavpaqlgfvhcepdckalcmppspigcelvkepgcgccmcalaegqac 66
 Qy 14 WLLLCSPAVGQNCSPGCRCPDEPAPRCR-PA--GVSLVLD-GGCCRVCARQLGELC 68

Db 67 gvytere-ag-qlrc 79
 Qy 69 TE-RDPCDPKHLGFC 82

RESULT 12
 ID IBP4 HUMAN STANDARD; PRT; 258 AA.
 AC P22632;
 DT 01-AUG-1991 (REL. 19, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 PRECURSOR (IGFBP-4)
 DE (IBP-4) (IGF-BINDING PROTEIN 4).
 GN IGFBP4 OR IBP4.

OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91186988.
 RA LATOUR D., MOHAN S., LINKHART T.A., BAYLINK D.J., STRONG D.D.;
 RL MOL. ENDOCRINOL. 4:1806-1814(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE; 91133415.

RA SHIMASAKI S., UCHIYAMA F., SHIMONAKA M., LING N.;
 RL MOL. ENDOCRINOL. 4:1451-1458(1990).
 RN [3]
 RP SEQUENCE FROM N.A.; AND SEQUENCE OF 22-41.

RC TISSUE=OSTEOSARCOMA;
 RX MEDLINE; 91223006.
 RA KIEFER M.C., WASTARZ F.R., BAUER D.M., ZAPF J.;
 RL J. BIOL. CHEM. 266:9043-9049(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;

RA STRONG D.D., MORALES S., LEE K., BOONYARATANAKORNKIT V.,
 RA BAYLINK D.J., MOHAN S.;
 RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [5]
 RP SEQUENCE OF 22-53.

RC TISSUE=COLON;
 RX MEDLINE; 91235178.
 RA CULOUSCOU J.-M., SHOYAB M.;
 RL CANCER RES. 51:2813-2819(1991).
 CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
 CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.

CC -!- BINDS IGF-II MORE THAN IGF-I.
 CC -!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
 DR EMBL; M38177; -; NOT ANNOTATED_CDS.
 DR EMBL; M62403; G184816; -.
 DR EMBL; U20982; G695254; -.
 DR PIR; A36549; A36549.
 DR PIR; B37252; B37252.
 DR PIR; B39842; B39842.
 DR MIM; 146733; -.

DR PROSITE; PS00222; IGF BINDING.
 DR PROSITE; PS00484; THYROGLOBULIN_1.
 KW GROWTH FACTOR BINDING; SIGNAL; GLYCOPROTEIN.
 FT SIGNAL 1 21
 FT CHAIN 22 258 INSULIN-LIKE GROWTH FACTOR BINDING
 FT CHAIN 22 258 POTENTIAL.
 FT CARBOHYD 125 125 THYROGLOBULIN TYPE 1.
 FT DOMAIN 200 249 P-> A (IN REF. 1, 4 AND 5).
 FT CONFLICT 51 51 I -> F (IN REF. 1 AND 4).
 FT CONFLICT 198 198
 SQ SEQUENCE 258 AA; 27934 MW; 58AC8AC3 CRC32;

Query Match 4.5%; Score 123; DB 5; Length 258;
 Best Local Similarity 38.2%; Pred. No. 4.54e-07;
 Matches 26; Conservative 14; Mismatches 22; Indels 6; Gaps 6;

Db 1 mlplclvaalllaagpgpslgde-aihcpcpseecklarcpvpgceelvrepqgcac 59
 Qy 6 MGPVRFVAFVLLALCSRPVAVGQNCSPGCR-CPDEPAPRCR-PA-GVS-LVLD-GGCCRVC 60

Db 60 alglmpc 67
 Qy 61 AKQLGELC 68

RESULT 13
 ID 60A DROME STANDARD; PRT; 455 AA.
 AC P27091;
 DT 01-AUG-1992 (REL. 23, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE 60A PROTEIN PRECURSOR.
 GN 60A OR TGFβ-60A.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92021021.

WHARTON K.A., THOMSEN G.H., CELBART W.M.;
PROC. NATL. ACAD. SCI. U.S.A. 88:9214-9218(1991).
(2)
SEQUENCE FROM N.A.
MEDLINE; 92290120.
DOCTOR J.S., JACKSON P.D., RASHKA K.E., VISALLI M., HOFFMANN F.M.;
DEV. BIOL. 151:491-505(1992).
-!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH PEAKS
OF TRANSCRIPTION DURING EARLY EMBRYOGENESIS, IN PUPAE, AND IN
ADULT MALES.
-1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
-!- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.
EMBL; M77012; G156728; -;
EMBL; M84795; G156730; -;
PIR; A41233; A41233.
PIR; A43918; A43918.
HSSP; P08112; 1TFG.
DR FLYBASE; FBGN004788; TGF-BETA-60A.
PROSITE; PS00250; TGF BETA.
GROWTH FACTOR, CYTOKINE; GLYCOPROTEIN; SIGNAL.
SIGNAL ?
FT PROPEP 1 ?
FT PROPEP 335 POTENTIAL.
FT CHAIN 336 455 60A PROTEIN.
FT DISULFD 354 420 BY SIMILARITY.
FT DISULFD 383 452 BY SIMILARITY.
FT DISULFD 387 454 BY SIMILARITY.
FT DISULFD 419 419 INTERCHAIN (BY SIMILARITY) .
FT CARBOHYD 238 238 POTENTIAL.
FT CARBOHYD 250 250 POTENTIAL.
FT CARBOHYD 396 396 POTENTIAL.
SQ SEQUENCE 455 AA; 51687 MW; 6FRAC7CD CRC32;

Query Match 4.5%; Score 121; DB 1; Length 455;
Best Local Similarity 43.2%; Pred.No.1.09e-06;
Matches 16; Conservative 9; Mismatches 10; Indels 2; Gaps

Ddb 419 ccaprtgialpvlyhlndenvnlkkyrmnvkacgc 455
||| | : ||| | : | : ||| | : ||| : |||
Qy 292 CCTHRITTLVEFKCPDGEV-MKKQM-MFIKTCACH 326

RESULT 14
ID IBP4 MOUSE STANDARD; PRT; 254 AA.
AC P47879;
DT 01-FEB-1996 (REL. 33, CREATED)
DD 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DE 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 PRECURSOR (IGFBP-4)
(IBP-4) (IGF-BINDING PROTEIN 4).
GN IGFBP4 OR IGFBP-4
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC EUTHERIA; RODENTIA.
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 95121750.
RX SCHULLER A.G.P., GROFFEN C., VAN NECK J.W., ZWARTHOF E.C.,
AL DROP S.L.S.;
AL MOL. CELL. ENDOCRINOL. 104:57-66(1994).

```
[2]
RN      SEQUENCE FROM N.A.
RC     STRAIN=C57BL/6J;
RA     MEDLINE; 95152444.
RX     BETHEL C.R., VITULLO J.C., MILLER R.F., ARON D.C.;
RL     BIOCHEM. MOL. BIOL. INT. 34:385-392(1994).
CC     -I- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFS
CC     AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
CC     PROMOTING EFFECTS OF THE IGFS ON CELL CULTURE. THEY ALTER THE
CC     INTERACTION OF IGFS WITH THEIR CELL SURFACE RECEPTORS.
CC     -I- BINDS IGF-II MORE THAN IGF-I
CC     -I- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
DR     EMBL; X81582; G550383; -.
DR     EMBL; X76066; G416033; -.
RW     GROWTH FACTOR BINDING; SIGNAL; GLYCOPROTEIN.
FT     SIGNAL   1    21
          FT     22    254
          FT
          FT     DOMAIN        196    245
          FT     CARBOHYD     125    125
          FT     CONFLICT     4       5
          FT     CONFLICT    13      13
          FT     CONFLICT    56      56
          FT     CONFLICT    67      67
          FT     CONFLICT    67      67
          FT     SEQUENCE     254 AA; 27760 MW; 7C9BC4EE CRC32;
SQ
Query Match              4.3%; Score 116; DB 5; Length 254;
Best Local Similarity 36.9%; Pred.No. 9.35e+06;
Matches 24; Conservative 14; Mismatches 21; Indels 6; Gaps 6;
Ddb      1 mlpcslvaallltagrpalgde-aihpccseeeklarcrppvgceelvrepqcgscatc 59
Qy       | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
         | : : : : : : : : : : : : : : : : : : : : : : : : : : : |
         6 MGPRVAFVVLLALCRPAVGQCNGPCR-CFDEPAPRC-PA-GVS-LVID-GCGGCRVC 60
Db       60 algiq 64
Qy       |||
         61 AKQLG 65
RESULT 15
RID IBP1 RAT STANDARD; PRT; 272 AA.
DAC P21743;
RAC 01-MAY-1991 (REL. 18, CREATED)
DTT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DDT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DDE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (IGFBP-1)
DDE (IBP-1) (IGF-BINDING PROTEIN 1).
DGN IGFBP1 OR IGFBP-1.
DNC RATTUS NORVEGICUS (RAT).
DSC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUOTHERIA; RODENTIA.
[1]
SEQUENCE FROM N.A.
RP RP
ARC TISSUE=DECIDUA;
RAC MEDLINE; 90231347.
SA MURPHY L.J., SENEVIRATNE C., BALLEJO G., CROZE F., KENNEDY T.G.;
RL MOL. ENDOCRINOL. 4:329-336(1990).
[2]
SEQUENCE FROM N.A.
RC TISSUE=LIVER;
```

```

RX MEDLINE; 91141487.
RA MOHN K.L., MELBY A.E., TEWARI D.S., LAZ T.M., TAUB R.A.;
RL MOL. CELL. BIOL. 11:1393-1401(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93149132.
RA OOI G.T., TSENG L.Y.H., TRAN M.Q., RECHLER M.M.;
RL MOL. ENDOCRINOL. 6:2219-2228(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX MEDLINE; 94250701.
RA LACSON R.G., OEHLER D., YANG E., GOSWAMI R., UNTERMAN T.G.;
RL BIOCHIM. BIOPHYS. ACTA 1218:95-98(1994).
RN [5]
RP SEQUENCE OF 26-59.
RX MEDLINE; 90322923.
RA UNTERMAN T.G., OEHLER D.T., GOTWAY M.B., MORRIS P.W.;
RL ENDOCRINOLOGY 127:789-797(1990).
CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
CC -1- BINDS EQUALLY WELL IGF-I AND IGF-II.
CC -1- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
DR EMBL; M58634; G204733; -.
DR EMBL; M89791; G204737; -.
DR EMBL; L22979; G1098473; -.
DR FIR; A36082; A36082.
DR FIR; A37398; A37398.
DR FIR; A39683; A39683.
DR PROSITE; PS00222; IGF BINDING.
DR PROSITE; PS00484; THYROGLOBULIN_1.
KW GROWTH FACTOR BINDING; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 272
FT INSULIN-LIKE GROWTH FACTOR BINDING
FT PROTEIN 1.
FT DOMAIN 215 264
FT SITE 259 261
FT CELL ATTACHMENT SITE.
FT R -> A (IN REF. 1 AND 4).
FT CONFLICT 79 79
FT CONFLICT 111 112
FT CONFLICT 201 201
FT CONFLICT 265 265
FT H -> N (IN REF. 1).
SQ SEQUENCE 272 AA; 29684 MW; 59EDF790 CRC32;

Query Match 4.2%; Score 113; DB 5; Length 272;
Best Local Similarity 42.9%; Pred. No. 3.32e-05;
Matches 21; Conservative 4; Mismatches 18; Indels 6; Gaps 6;

Db 43 cppvpaas-cpe-lsrp-agcgccptcalplgaacgvatarc-aq-glsc 86
QY 35 CPDEFAPRCFAGVSLVLDGCGCCRCVCAKQGLGELC-TERDPCDPHKLFC 82

```

Search completed: Wed Sep 17 09:28:33 1997
Job time : 51 secs.